

**Gezauberte Ereignisse als Zugang zur kognitiven
Entwicklung: Kindliche Vorstellungen über
Eigenschaften intensiver Grössen**

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der Universität Zürich

vorgelegt von
Priska Simone Schmid Haller
von
Buchs/ZH, Zürich und Reinach/AG

Angenommen auf Antrag von
Herrn Prof. Dr. Friedrich Wilkening
und
Herrn Prof. Dr. Wolfgang Marx

Studentendruckerei, Zürich, 2004

Gezauberte Ereignisse als Zugang zur kognitiven Entwicklung: Kindliche Vorstellungen über Eigenschaften intensiver Grössen

Mit Mischungsaufgaben von sogenannten intensiven Grössen kann der Erwerb von physikalischem Regelwissen untersucht werden. Intensive Grössen wie etwa die Farbhellickeit von Flüssigkeiten sind physikalische Grössen, bei denen eine Zusammenfü-gung zum Durchschnitt der Grössen führt. Untersuchungen mit Mischungsaufgaben von intensiven Grössen zeigen, dass viele Grundschulkinder fehlerhafte Intensivierungs- bzw. Additionsvorstellungen haben, anstatt die korrekte Durchschnittsregel anzuwen-den. Mit Hilfe einer neuartigen Methode, dem *magic experiment*, wurde eine Erfassung von erfahrungsnahem Regelwissen über Mischungsvorgänge von gleichfarbigen Flüs-sigkeiten ermöglicht. Kindern zwischen 6 und 12 Jahren sowie Erwachsenen (N = 140) wurden physikalisch korrekte und unkorrekte Mischungsvorgänge präsentiert. Korrekte Mischungen entsprachen der Durchschnittsregel und unkorrekte, die mittels chemi-schem Trick dunkler waren als die Ausgangsflüssigkeiten, entsprachen der Additionsre-gel. Diese Mischungen mussten anhand ihres Realitätscharakters eingeschätzt werden, um die Anwendung von Durchschnitts- oder Additionsregeln zu bestimmen. Durch den Vergleich mit der Regelanwendung einer vorher gestellten Prädiktionsaufgabe konnte gezeigt werden, dass die Anwendung der Durchschnittsregel in beiden Aufgabentypen sehr stabil und die der Additionsregel viel instabiler war. Das *magic experiment* diente ferner der Erforschung des kindlichen Verständnisses über mögliche Verursacher der dargebotenen korrekten und unkorrekten Mischungen. Es zeigte sich bei jüngeren Kin-dern ein starker Glaube in die maximale Befähigung der Figur des Zauberers (magi-sches Denken), der ab dem zehnten Lebensjahr deutlich abnahm.

Magical events as a tool for studying children's ideas about intensive quantities

Mixture tasks of intensive quantities allow the study of children's knowledge acquisition about physical laws. If two different intensive quantities e.g. two liquids with different color intensities are mixed together, the resulting intensity of the mixture formed is the average of the two initial intensities. Previous studies using mixture tasks indicate that many children tend to rely on inadequate adding rules, instead of using the correct averaging rule. A new method, the "magic experiment", is introduced. It enables an intuitive assessment of the rule knowledge in a task of mixing two liquids with different color intensities. In the magic experiment, participants were presented with physically correct and incorrect mixtures. Correct mixtures corresponded to the averaging rule and incorrect ones, which were darker than the two initial components (attained with the help of a chemical trick), corresponded to the adding rule. Participants had to judge the different mixtures by their status of reality and these judgments were assigned to the use of the averaging or adding rule. Children aged 6 to 12 years and adults ($N = 140$) performed a standard mixture task followed by the magic experiment. Comparing rule use in the magic experiment to rule use in the standard task showed that the use of the averaging rule was very stable across tasks from 8 years of age upwards, whereas the use of the adding rule was more unstable and task dependent throughout childhood. The magic experiment also examined children's ideas about various agents who could produce the presented trick. Results showed a strong belief in the capability of the character of a magician (i.e. magical thinking) at the age of 6 and 8 years that decreased strongly from the age of 10 years on.

1	Allgemeine Einleitung	1
2	Theorie	5
2.1	Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen	5
2.1.1	Bisherige Befunde zur kognitiven Durchschnittsbildung bei intensiven Grössen	6
2.1.2	Bisherige theoretische Konzeptionen zur Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen	17
2.2	Theorie des <i>extinction paradigm</i> bzw. <i>magic experiment</i>	28
2.2.1	Studien zum <i>extinction paradigm</i>	28
2.2.2	Ausweitung des Anwendungsbereichs: <i>Magic experiment</i>	33
2.2.3	Kritik an der Methode des <i>extinction paradigm</i> bzw. <i>magic experiment</i>	35
2.3	Kindliches Verständnis von gezauberten Ereignissen	39
2.3.1	Unterscheidungsfähigkeit zwischen gezauberten und nicht-gezauberten Ereignissen bei Vorschulkindern	39
2.3.2	Entwicklung des Verständnisses der Verursachung von gezauberten Ereignissen	43
3	Empirische Untersuchung	50
3.1	Methode	50
3.1.1	Versuchspersonen	52
3.1.2	Versuchsmaterial	53
3.1.3	Versuchsdurchführung	58
3.1.4	Versuchsplan	59
3.2	Prädiktion der Farbmischungsergebnisse	62
3.2.1	Fragestellung	62
3.2.2	Methode	63
3.2.3	Auswertung	64
3.2.4	Ergebnisse	67
3.2.5	Diskussion	81
3.3	Magic experiment	92
3.3.1	Fragestellung	92
3.3.2	Methode	99
3.3.3	Auswertung	101

3.3.4	Ergebnisse.....	105
3.3.5	Diskussion	151
3.4	Kausalitätsverständnis gezauberter Ereignisse.....	168
3.4.1	Fragestellung	168
3.4.2	Methode	170
3.4.3	Auswertung.....	172
3.4.4	Ergebnisse.....	173
3.4.5	Diskussion	215
4	Allgemeine Diskussion.....	225
4.1	Theoretische Schlussfolgerungen zur Entwicklung der kognitiven Durch-	
	schnittsbildung bei intensiven Grössen	225
4.1.1	Adaptive Problemlösung bei der Mischung intensiver Grössen	227
4.1.2	Gewissheit und Stabilität der verschiedenen Lösungsstrategien im	
	Entwicklungsverlauf.....	231
4.1.3	Ausblick auf weiterführende Untersuchungen	239
4.2	Kausalitätsverständnis gezauberter Ereignisse.....	242
5	Literaturverzeichnis	255
	Anhang A: Instruktion des Experimentes.....	268
	Anhang B: Farbintensitätsskala	275
	Anhang C: Darstellung der einzelnen Versuchsdurchgänge	276
	Anhang D: Rating-Skalen.....	281

1 Allgemeine Einleitung

Ein wesentlicher Teil der Psychologie der kognitiven Entwicklung beschäftigt sich mit der Entstehung und Entwicklung naiver Vorstellungen zu Begriffen, Phänomenen und Prinzipien, die sich die Kinder auf der Basis vielfältiger Alltagserfahrungen aneignen. Ebenso zentraler Gegenstand der kognitiven Entwicklungspsychologie ist die Umwandlung dieser vorunterrichtlichen Alltagsvorstellungen (*common sense*) in wissenschaftliche Vorstellungen.

In der Vergangenheit hat sich die Domäne der intuitiven Physik als äusserst geeignetes Feld zur Erforschung der Entwicklung kindlicher Vorstellungen erwiesen (vgl. Anderson & Wilkening, 1991). Physikalische Naturgesetze laufen immer gleich ab und werden bei der Interaktion mit der Natur stets korrekt zurückgemeldet, wodurch der Wissensaufbau erheblich erleichtert wird. Man weiss heute, dass physikalisches Wissen schon früh auf verschiedenen Ebenen wie der Motorik (z. B. *embodied knowledge*), der Wahrnehmung und der Kognition repräsentiert ist. Die Erforschung der Zusammenhänge der Repräsentationen auf diesen verschiedenen Ebenen ist heute Gegenstand vieler Untersuchungen.

Ein weiteres Teilgebiet innerhalb des Bereichs der intuitiven Physik bilden die fehlerhaften Vorstellungen (sog. *misconceptions*), die manchmal sogar bis ins Erwachsenenalter erhalten bleiben. Die Annahme, dass physikalische Fehlannahmen durch generelle Beschränkungen der kognitiven Fähigkeiten zu Stande kommen, war lange Zeit weit verbreitet. Heute weiss man aber, dass die fehlerhaften Vorstellungen bei Kindern nicht nur als Ausdruck altersspezifischer kognitiver Defizite verstanden werden dürfen. Gerade bei Kindern scheint das physikalische Wissen in hohem Ausmass in verschiedenen Wissenskomponenten auf unterschiedlichen Repräsentationsebenen gespeichert zu sein, die in Abhängigkeit von der Aufgabensituation durchaus zu korrekten Lösungen physikalischer Problemstellungen führen können (vgl. Wilkening & Lamsfuss, 1993).

In der vorliegenden Arbeit wurden hierzu kindliche Vorstellungen über ein physikalisches Prinzip untersucht, über deren Entwicklung und Repräsentationen man noch wenig weiss: Die kindlichen Vorstellungen über Eigenschaften intensiver Grössen.

Intensive Grössen bezeichnen physikalische Zustandsgrössen und werden gegenüber von extensiven Grössen abgegrenzt. Extensive Grössen machen eine Aussage über die Mengenartigkeit eines physikalischen Systems und geben darüber Auskunft, „wie viel“ von etwas vorhanden ist. Beispiele für extensive Quantitäten sind das Volumen oder die Teilchenzahl eines Systems. Intensive Grössen hingegen zeigen an, wie „stark“ oder eben wie „intensiv“ etwas ist. Beispiele hierfür sind die Temperatur oder die Farbintensität eines Systems. Die beiden Zustandsgrössen können auch durch ihr Verhalten beim Zusammenfügen definiert werden: Werden zwei extensive Grössen zusammengefügt, so addieren sich ihre Ausprägungen. Werden zwei intensive Grössen zusammengefügt, so ergibt sich immer eine mittlere Ausprägung.

Verschiedene entwicklungspsychologische Studien zur Vorstellung vom Zusammenfügen intensiver Grössen zeigten, dass viele Kinder hierzu eine fehlerhafte Vorstellung in dem Sinne haben, dass die intensiven Grössen addiert werden. Konkret waren die Kinder etwa der Ansicht, dass zwei gleich warme Flüssigkeiten nach dem Zusammengiesen wärmer, zwei gleiche hellrote Flüssigkeiten dunkler oder zwei gleich süsse Flüssigkeiten süsser würden. Auch hier stellte sich die Frage, was es schwierig macht, die richtigen Naturgesetze trotz vielfacher Erfahrung zu lernen.

Wichtige Hinweise zur Aufklärung dieser fehlerhaften Vorstellung kamen aus informationsintegrationstheoretischen Studien. Es konnte nachgewiesen werden, dass beim Zusammenfügen intensiver Grössen die relevanten Grössen Informationen tatsächlich miteinander verknüpft werden (Jäger, 1996; Jäger & Wilkening, 2001). Bei der fehlerhaften Vorstellung einer Intensivierung nach dem Zusammenfügen werden die intensiven Grössen allerdings addiert, anstatt dass der Durchschnitt gebildet wird. Die korrekte Verknüpfungsregel beim Zusammenfügen intensiver Grössen entspricht in der Terminologie der Informationsintegrationstheorie der kognitiven Durchschnittsbildung (*cognitive averaging*). Aus Studien aus dem sozialen Bereich konnte nachgewiesen werden, dass Kinder an sich schon früh zur Durchschnittsbildung befähigt sind (z. B. Butzin & Anderson, 1973). Die Anwendung vereinfachender Lösungsregeln bei der Integration intensiver Grössen kann deshalb nicht Ausdruck grundsätzlicher altersspezifischer Beschränkungen verstanden werden.

Ein weiteres Ergebnis früherer Untersuchungen besteht im Nachweis der hohen Abhängigkeit der kindlichen Problemlösungsstrategien von der Aufgabendarbietung. Der Einsatz von korrekten und fehlerhaften Lösungsvorstellungen hängt also auch beim Zusammenfügen von intensiven Grössen nicht nur von der Verfügbarkeit der Regeln ab, sondern auch von den als lösungsrelevant erachteten Situationsmerkmalen der Aufgaben.

Eine direkte Überprüfung und Identifizierung von relevanten Aufgabeneinflüssen mithilfe experimenteller Untersuchungen stand allerdings bisher noch aus. In der vorliegenden Arbeit wurde dieses Vorhaben am Beispiel der intensiven Grösse Farbintensität erstmals umgesetzt. Gleichzeitig wurde eine neue Methode eingeführt, die eine Erforschung der Stabilität und Reife von Wissen (*in-the-system-knowledge*) ermöglicht: der Präsentation gezauberter und nicht-gezauberter Ereignisse und der Erfassung der Reaktionen auf diese Ereignisse. Die neue Methode kann als starke Modifikation bereits bekannter Methoden wie dem *extinction paradigm* (z. B. Miller, 1976), dem *magic experiment* (z. B. Lindsay & Creedon, 1985) oder auch der bei vorsprachlichen Kindern eingesetzten *violation of expectancy* (z. B. Baillargeon, 1986, 1993, 1996) betrachtet werden. Reagiert das Kind auf ein gezaubertes Ereignis mit einer Erhaltung des dabei verletzten physikalischen Prinzips, so ist sein physikalisches Wissen auf einer reiferen Stufe einzuordnen, als wenn das Prinzip als Reaktion auf das Gesehene verworfen wird.

In einer ersten Phase wurden in einer neuen Aufgabensituation zur Mischung von farbigen Flüssigkeiten die prädiktiven Lösungsregeln bei insgesamt 120 Kindern im Alter zwischen sechs und zwölf Jahren sowie bei 20 Erwachsenen identifiziert. Die experimentell variierte Aufgabensituation sollte eine Identifizierung lösungsrelevanter Hinweise der Aufgabendarbietungen zulassen. Im Anschluss und zu Beginn des zweiten Teils wurden die Komponenten von einer Zauberin tatsächlich gemischt und durch den Einsatz von chemischen Tricks wurden verschiedene Mischungsergebnisse zurückgemeldet: Physikalisch korrekte, der Durchschnittsregel entsprechende sowie physikalisch unkorrekte, der additiven Mischungsvorstellung entsprechende Mischungsergebnisse. Die Flüssigkeiten enthielten das eine Mal angefärbtes Wasser und das andere Mal waren sie zusätzlich mit unsichtbarem chemischen Zusatz versetzt, so dass sich durch das Zusammengiessen derselben Ausgangsgläser unbemerkt je eine normale Mischung und eine dunklere Mischung ergaben.

Im zweiten Teil wurden die Reaktionen auf die verschiedenen Mischungsresultate mit zwei neuartigen abhängigen Variablen erfasst, die eine Aussage über die Stabilität sowie über den Grad der Gewissheit bzw. der Reife der Mischungsregeln aus Phase 1 ermöglichten. Durch die stark modifizierte Methode des *magic experiment* waren nicht nur Rückschlüsse auf die Entwicklung der Stabilität und Reife der korrekten Durchschnittsbildung möglich, sondern erstmals auch auf diejenige der fehlerhaften Intensivierungsvorstellung.

Im dritten Teil der vorliegenden Untersuchung wurde das gleiche Paradigma für das inhaltliche Verständnis von gezauberten Ereignissen der Kinder am Beispiel von Mischungen intensiver Grössen verwendet. Die präsentierten gezauberten und nicht gezauberten Ereignisse konnten als indirekter Zugang zum Kausalitätsverständnis gezauberter Ereignisse verwendet werden, indem die verschiedenen Ereignisse anhand ihrer Machbarkeit durch verschieden befähigte Akteure eingeschätzt werden mussten. Auf diese Art und Weise konnte mit der modifizierten Methode des *magic experiment* erstmals ein Einblick in das kindliche Verständnis über die Entstehung von gezauberten Ereignissen gewonnen werden, ohne auf die für Kinder schwierig zu artikulierenden Erklärungsversuche zugreifen zu müssen.

2 Theorie

2.1 Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen

Eine der vielfältigen Möglichkeiten, die Fülle von Umweltinformationen zu reduzieren, besteht in der Bildung des Durchschnitts von Merkmalen innerhalb einer Dimension. Die durchschnittliche Grösse eines Apfels hängt von der Beobachtung von verschiedenen Äpfeln ab und wird durch die mittlere Grösse bestimmt. Die durchschnittliche Leistung eines Fussballspielers schätzen wir aufgrund seiner Einzelleistungen über eine bestimmte Zeitspanne hinweg ein. Oder wir bestimmen – bereits etwas mathematischer – eine durchschnittliche Note durch die Gesamtzahl der Noten in Bezug auf die Anzahl der erbrachten Leistungsnachweise. Eine der wichtigsten Eigenschaften der Durchschnittsbildung besteht darin, dass der Durchschnitt immer zwischen den Extremwerten und nie darüber oder darunter liegt (*range principle*). Personen, die über solches Wissen in diesen Alltagsbeispielen verfügen, besitzen die Fähigkeit zur Operation der Durchschnittsbildung.

Wie wir aus Untersuchungen aus der intuitiven Physik wissen (z. B. Krist, Fieberg & Wilkening, 1993), werden algebraische Operationen von Kindern nicht erst durch die Vermittlung in der Schule auf der formalen Ebene angeeignet, sondern sie werden im Alltag durchaus auch schon früher korrekt angewendet. Zu welchem Zeitpunkt in der Entwicklung wird diese Fähigkeit zur Durchschnittsbildung gebildet? Im Gegensatz zu anderen algebraischen Operationen wie der Addition und der Multiplikation (zur Addition z. B. Anderson & Cuneo, 1978a, 1978b; zur Multiplikation z. B. Wilkening, 1981, 1982), weiss man zur kognitiven Entwicklung der Durchschnittsbildung allerdings noch nicht viel.

Eine Studie aus dem Bereich der sozialen Kognition konnte zeigen, dass schon fünfjährige Kinder zur Integration verschiedener Informationsteile mittels Durchschnittsbildung bei der Einschätzung der Beliebtheit von Spielzeugen befähigt waren (Butzin & Anderson, 1973). Anders scheint es allerdings im physikalischen Bereich zu sein: Di-

verse Studien zur kognitiven Durchschnittsbildung bei intensiven Grössen zeigten bei Kindern einheitlich erhebliche Schwierigkeiten in der korrekten Anwendung dieser Regel. Es zeigten sich nicht nur grobe Verletzungen des scheinbar so einfachen *range principle*, sondern auch die Verwendung von Regeln, die in anderen Anwendungsbereichen durchaus korrekt sind, im Zusammenhang mit intensiven Grössen aber zu frapierend unkorrekten Lösungsversuchen führen.

Im folgenden Kapitel werden die wichtigsten Untersuchungen zur Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen beschrieben. Aufgrund der widersprüchlichen Befunde und entsprechend uneinheitlicher Erklärungsversuche erschien es nötig, diese detailliert darzustellen, um die heute plausibelste Erklärung mit dem Ansatz des adaptiven Denkens verständlich zu machen. Gleichzeitig veranschaulichen die Studien und Theorien aus gut zwei Jahrzehnten eindrücklich den Prozess der Erkenntnisgewinnung durch die Wechselwirkung von Theorie und Empirie.

2.1.1 Bisherige Befunde zur kognitiven Durchschnittsbildung bei intensiven Grössen

Untersuchungen der Forschergruppe um Sidney Strauss in den achtziger Jahren

Bei den Studien von Strauss und Mitarbeitern standen Untersuchungen zu den intensiven Grössen Süsse und Temperatur im Vordergrund (z. B. Stavy, 1981; Stavy, Strauss, Orpaz & Carmi, 1982; Strauss & Stavy, 1982). An dieser Stelle wird die Aufgabe zur Süsse (Stavy et al., 1982) exemplarisch für die Studien zur Mischung von intensiven Grössen der Forschergruppe um Sidney Strauss vorgestellt.

Kindern im Alter von vier bis dreizehn Jahren wurden drei zur Hälfte mit Wasser gefüllte Gläser (A, B, C) präsentiert. In jedes dieser Gläser wurde vor den Augen der Kinder je ein Teelöffel mit Zucker zugegeben. Durch anschliessendes Umrühren wurde der Zucker im Wasser verteilt und aufgelöst. Nachdem die Kinder bestätigt hatten, dass alle drei Flüssigkeiten gleich süss waren, wurden die Flüssigkeiten aus den Gläsern A und B in ein viertes, leeres Glas zusammengeleert. Die Kinder hatten diese Mischungsflüssigkeit nun mit der Flüssigkeit in Glas C hinsichtlich ihrer Süsse zu vergleichen (gleich süss, süsser oder weniger süss?) und ihre Urteile zu begründen. Während die Vorschul-

kinder (Vier- und Fünfjährige) die Aufgabe vorwiegend korrekt lösten, fiel die Anzahl richtiger Antworten bei den Grundschulkindern ab und erreichte erst bei den Elf- bis Zwölfjährigen wieder ihre ursprüngliche Höhe. Besonders auffällig war, dass die jüngeren Grundschulkindern – die Achtjährigen sogar zu 90 Prozent – glaubten, die Mischungsflüssigkeit sei süsser als die Flüssigkeit in Glas C. Die Häufigkeiten richtiger Antworten zeigten im Entwicklungsverlauf der Spanne zwischen viertem und zwölftem Lebensjahr demnach einen u-förmigen Verlauf. Bei der Mischung von Flüssigkeiten mit der intensiven Grösse Temperatur, welche durch eine bestimmte, bei beiden Flüssigkeiten gleiche Anzahl von Kerzen symbolisiert wurde, stellte sich ein vergleichbares Ergebnis ein (Strauss & Stavy, 1982).

Anhand der Anzahl korrekter Antworten und den entsprechenden verbalen Begründungen der richtigen Antworten machten die Autoren drei Entwicklungsphasen beim Erwerb der richtigen Lösungsstrategie aus: Während die jüngeren Kinder ein intuitives erfahrungsgestütztes Verständnis der Aufgabe zur richtigen Problemlösung führte, dominierten bei den älteren Kindern zunehmend bidimensionale und proportionale Begründungen, die ebenfalls zur richtigen Lösung führten. Bei den jüngeren Grundschulkindern hingegen lag eine falsche, additive Problemlösung vor. Strauss und Mitarbeiter bezeichneten die frühe Phase des Grundschulalters als Übergangsphase, in der das Alltagswissen durch den zunehmenden Erwerb quantitativer Beziehungen durch die Schule überlagert war, was zu vorwiegend falschen Antworten geführt habe.

Die theoretische Konzeption von Strauss und Mitarbeitern auf der Basis dieser Befunde wird im Kapitel „Theorie der u-förmigen Entwicklung“ (S. 17ff.) ausführlicher dargelegt.

Untersuchungen um James Dixon und Colleen Moore in den neunziger Jahren

Dixon, Moore und weitere Mitarbeiter (Ahl, Moore & Dixon, 1992; Dixon & Moore, 1996; Moore, Dixon & Haines, 1992) untersuchten den Entwicklungsprozess der kognitiven Durchschnittsbildung anhand von Temperaturmischungen. Im Zentrum ihres Interesses stand die Beziehung zwischen intuitivem Verständnis bei verbalen Aufgaben und der Erzeugung von mathematischen Strategien zur Problemlösung in numerischen Aufgaben. Am Beispiel von Aufgaben zur Temperaturmischung wurden verschiedene Teilstrategien, die alle für die korrekte Durchschnittsbildung und Lösung der Aufgabe vorhanden sein müssen, untersucht. Durch die Erfassung der verwendeten Teilstrategien

von Kindern in verschiedenem Alter sollte ein Aufschluss über die Entwicklung des Verständnisses der Temperaturmischung möglich werden.

Dixon und Moore legten vier Prinzipien fest, die alle für ein perfektes Verständnis nötig sind und gleichzeitig ein reifes kognitives Durchschnittschema charakterisieren:

Beim ersten Prinzip *Above-Below* richtet sich der Fokus nur auf die Veränderung der ursprünglichen Temperatur einer Komponente, welche sich immer in Richtung der zugegebenen Temperatur verändert. Je nach zugegebener Temperatur wird die resultierende Temperatur über oder unter der ursprünglichen Temperatur der fokussierten Komponente liegen.

Das *Range Principle* besagt, dass die resultierende Temperatur der Mischung zweier Komponenten innerhalb des Bereichs der beiden ursprünglichen Temperaturen liegen muss.

Das *Crossover* Prinzip bezieht sich auf die Interaktion von intensiver und extensiver Quantität – am Beispiel der Temperaturmischung also von Temperatur und Volumen der Flüssigkeit. Je grösser das Volumen einer der beiden Komponenten ist, umso grösser ist der Effekt der Temperatur dieser Komponente auf die resultierende Temperatur der Mischungsflüssigkeit. Werden 2L einer Flüssigkeit mit 30 Grad statt 1L zugegeben, so ist der Effekt der Temperatur von 30 Grad grösser, als wenn nur 1L zugegeben wird. Die Richtung des Effektes hängt aber davon ab, ob die Temperatur dieser Komponente über oder unter derjenigen der anderen Komponente ist. Die Richtung des Effektes der obigen Flüssigkeit mit 30 Grad Temperatur hängt also davon ab, ob die andere Komponente wärmer oder kälter als 30 Grad ist. Ist sie wärmer, dann resultiert eine Mischungstemperatur über 30 Grad, wenn sie aber kälter ist, resultiert eine Mischung mit niedrigerer Temperatur.

Das *Equal-Intensities-Equal Principle* ist auf die Mischung von Komponenten fokussiert, deren intensiven Grössen gleich ausgeprägt sind. Wenn zwei gleich warme Komponenten gemischt werden, findet kein Temperaturaustausch statt, unabhängig vom Volumen der beiden Komponenten.

Die Ergebnisse der Studien von Dixon, Moore und Mitarbeitern zeigten, dass die Aufgaben der numerischen Bedingung, bei der die Temperaturinformation durch Zahlen vermittelt wurde, stets schlechter gelöst wurden als bei verbaler Präsentation der Tem-

peraturinformation, bei der die Temperatur verbal beschrieben (sehr warm, warm, mittel, etc.) und die Höhe der Temperatur auf einem Thermometer angezeigt wurden. Bei den jüngsten Kindern ihrer Studien, den Achtjährigen, zeigte sich in beiden Bedingungen vorwiegend kein Verständnis der Aufgaben. Aber auch bei den Zehn- bis Vierzehnjährigen wurde in hohem Ausmass auch bei der verbalen „intuitiven“ Bedingung noch mindestens eines dieser Prinzipien verletzt. Darunter waren auch nicht selten die scheinbar einfacheren Prinzipien *Range* und *Equal-Intensities-Equal*. Konkret gehörten hierzu Voraussagen wie „das Mischen von zwei Flüssigkeiten mit verschiedenen Temperaturen führt zu einer resultierenden Temperatur ausserhalb des Bereiches der zwei ursprünglichen“ oder „beim Mischen von zwei Flüssigkeiten mit der gleichen Temperatur verändert sich die resultierende Temperatur“. Ein annähernd vollständiges Verständnis von allen Prinzipien wurde erst in den Gruppen der 11. Klasse und der Erwachsenen gefunden.

Die Schlussfolgerungen der Befunde von Dixon, Moore und Mitarbeitern resultierten in einem Modell zur Entwicklung der Durchschnittsbildung mit zwei gleichzeitig ablaufenden Prozessen: Einerseits entwickelt sich die Fähigkeit, sich der Regeln bewusst zu werden. Andererseits wird der Wechsel zwischen intuitivem (alltagsnahes Wissen; führt bei intuitiven Aufgaben zur richtigen Lösung) und analytischem Denken (Schulwissen; Anwendung von schulgelernten Regeln) zunehmend leichter. Auch diese theoretische Konzeption wird im Verlauf dieser Arbeit (S. 23ff.) noch detaillierter beschrieben.

Untersuchungen von Stephan Jäger in den neunziger Jahren

In den neunziger Jahren führte die Gruppe um Jäger Untersuchungen mit der intensiven Grösse Farbhelligkeit durch und konnte damit die Erkenntnis zur Entwicklung der Durchschnittsbildung entscheidend vorantreiben. Die Aufgaben beinhalteten stets Mischungsvorgänge von zwei gleich farbigen, aber verschieden hellen Flüssigkeiten (Jäger, 1996, 1997; Jäger & Becker, 1997; Jäger & Wilkening, 2001).

Mit der Farbintensität oder Farbhelligkeit als intensive Grösse war es erstmals möglich, die relevante Information nicht mehr nur symbolisch sondern direkt visuell wahrnehmbar darzubieten. Eine Eigenheit der früheren Forschungen zur Durchschnittsbildung bei intensiven Quantitäten war, dass die relevante Information, vorwiegend Temperatur oder Süsse, immer nur in symbolisierter Form wie etwa mit Thermometer, Kerzen oder

Zuckerwürfel präsentiert wurde. Dadurch wurde die intensive Grösse auch zwangsläufig in numerischer Form wie z. B. durch die Anzahl von Kerzen oder Zuckerwürfeln dargeboten. Im Gegensatz zur extensiven Grösse in Form des Volumens der Flüssigkeit konnte die intensive Grösse vom Kind nicht direkt wahrgenommen werden. Die Verwendung von Farbe erlaubte nebst der direkten Wahrnehmbarkeit der dargebotenen Stimuli auch die Herstellung einer einfachen, direkten Beziehung zwischen der vorausgesagten Mischungsintensität und der Urteilsskala. Die vorausgesagten Mischungsintensitäten konnten in Form einer vielfach abgestuften Farbskala eingeschätzt werden, wodurch eine direktere und differenziertere Einschätzung der physikalischen Quantität ermöglicht werden sollte.

Ein weiteres Interesse von Jäger bestand im Nachweis von Verknüpfungsregeln bei der kognitiven Durchschnittsbildung. Hierzu wurden die Experimente so angelegt, dass überprüft werden konnte, ob die relevanten Grössen beim Lösungsprozess auch tatsächlich berücksichtigt und nach welcher Regel sie verknüpft wurden. Nebst dem Nachweis von Regeln wie der Durchschnitts- oder Additionsregel stand auch die Entwicklungsabfolge der Verwendung der nachgewiesenen Regeln im Vordergrund.

Experiment 1: Direkt wahrnehmbare Darbietung der intensiven Grösse und abgestufte Urteilsskala bei Mischungen verschiedener Intensitätskombinationen

In einer ersten Untersuchung führte Jäger (1996, Experiment 1) die neue intensive Grösse Farbintensität ein. Die Farbintensität brachte im Gegensatz zu den früheren Arbeiten von Strauss und Mitarbeitern (Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982) zwei entscheidende Neuerungen mit sich: Erstens wurden die dargebotenen Stimuli direkt visuell wahrnehmbar und nicht mehr symbolisch-numerisch dargestellt. Zweitens erfolgte die Einschätzung der vorausgesagten Mischung auf einer direkteren und abgestuften Farbskala und nicht mehr anhand der Wahlaufgabe bzw. Paarvergleichsmethode mit den drei Antwortmöglichkeiten „gleich geblieben“ „intensiver geworden“ (z. B. süsser oder wärmer) „weniger intensiv geworden“ (z. B. weniger süss oder weniger warm). Aufgrund der anschaulicheren Darbietung der relevanten Grösse wurde erwartet, dass ein stetig ansteigender und nicht mehr ein u-förmiger Entwicklungsverlauf resultiert. Bei Jägers Experiment sollten zudem nicht nur Paare mit der gleich ausgeprägten intensiven Grösse zusammengeleert werden, sondern auch Kombinationen mit unterschiedlich ausgeprägten intensiven Grössen. Dieser Anordnung lag

die Frage zu Grunde, ob es verschiedene Lösungsstrategien für die Aufgaben zur Mischung von gleichen und unterschiedlichen Farbintensitäten gibt.

In diesem Experiment hatten die Kinder (Sechs-, Acht-, Zehn- und Zwölfjährige) sowie Erwachsene die Mischungsergebnisse der Kombination gleicher Farbintensitäten (Helligkeiten 4 auf der Farbskala) und der Kombination unterschiedlicher Farbintensitäten (Helligkeiten 3 und 5 auf der Farbskala) anhand einer 7-teiligen abgestuften Farbskala einzuschätzen. Die Antworten wurden kategorisiert und verschiedenen Lösungsstrategien zugeordnet. Bei beiden Kombinationen waren hypothetisch die korrekte Durchschnittsregel (gemittelte Ausprägung beider Komponenten), die Additionsregel (dunkler als die dunklere Komponente) und die Subtraktionsregel (heller als die hellere Komponente) zu erwarten. Bei der Mischung unterschiedlicher Intensitäten war mit der Einschätzung auf der 7-teiligen Farbskala zudem auch die Anwendung einer Maximumregel (Zentrierung auf die dunklere Ursprungsintensität) und einer Minimumregel (Zentrierung auf die hellere Ursprungsintensität) möglich. Da aber nur ein Experimentaldurchgang pro Intensitätskombination durchgeführt wurde, konnte keine Aussage über die Konsistenz der Regelanwendung innerhalb einer Kombination gemacht werden.

Die Mischungsregeln wurden für die beiden Intensitätskombinationen getrennt ausgewertet, da sich die Entwicklungsverläufe der beiden häufigsten Mischungsregeln gemäss der Addition und der Durchschnittsbildung bei den beiden Bedingungen unterschieden hatten.

Bei Mischung gleicher Farbintensitäten zeigte sich eine erstaunlich häufige Anwendung der korrekten Durchschnittsregel. Bereits 70 Prozent der Sechsjährigen wussten, dass die Farbintensität der Mischungsflüssigkeit gleich bleibt. Der Anstieg mit zunehmendem Alter verlief nicht ganz linear, da die Zehnjährigen weniger häufig richtige Mischungsergebnisse voraussagten als die Achtjährigen. Nebst der Durchschnittsregel wurde nur noch die Additionsregel angewendet. Ein helleres Mischungsergebnis, was der Subtraktionsregel entsprochen hätte, wurde gar nie vorausgesagt. Die Häufigkeit der Anwendung der Additionsregel verlief demnach reziprok zur Häufigkeit der Durchschnittsbildung.

Die Ergebnisse bei der unterschiedlichen Intensitätskombination wiesen, mit Ausnahme der Altersgruppe der Zehnjährigen, auf zwei sehr deutliche gegenläufige Entwicklungstrends hin: Während die Anwendung der Additionsregel mit zunehmendem Alter abnahm, stieg die Anwendung der Durchschnittsregel an. Die meisten Sechsjährigen glaubten, das Mischungsergebnis werde dunkler als die dunklere der beiden Komponenten. Wie schon bei der Kombination gleicher Farbintensitäten lösten die Zehnjährigen die Aufgabe weniger häufig nach der korrekten Durchschnittsregel als die Achtjährigen. Die drei anderen möglichen Mischungsregeln wurden hingegen nur vereinzelt angewendet.

Bei der Interpretation seiner Resultate stützte sich Jäger zunächst auf den nicht gefundenen u-förmigen Verlauf der kognitiven Entwicklung bei der Mischung gleicher Farbintensitäten. Jäger machte den Unterschied zu den Ergebnissen von Strauss und Mitarbeitern (Strauss & Stavy, 1982; Strauss et al., 1982) an den unterschiedlichen Aufgabendarbietungen der beiden Methoden fest: Bei der Aufgabendarbietung von numerischen Temperaturangaben oder von Würfelzuckerstücken habe eine unreflektierte Addition stattgefunden, welche durch die direkte visuelle Darbietung der Farben verhindert worden sei. Mit der direkten visuellen Darbietung könnten die Kinder die relevante Dimension leichter erfassen und vermutlich werde dadurch auch erreicht, dass sie sich ein inneres Bild des Mischungsergebnisses machen können (Jäger, 1996, S. 73).

Als weiteres wichtiges Ergebnis bewertete Jäger die hohen Anteile der additiven Regel bei der Mischung unterschiedlicher Farbintensitäten. Als Erklärung für die Entstehung der Additionsvorstellung zog er sowohl die Volumen- als auch die Teilchentheorie herbei (Jäger, 1996, S. 74). Die Volumentheorie besagt, dass bei der Aufgabenlösung ein unreflektiertes Übertragen der Eigenschaft der extensiven Grösse auf die intensive Grösse stattfindet. Aus dem Anstieg des Volumens im Mischungsglas wird geschlossen, dass auch die Farbintensität zunehmen muss. Bei der Teilchentheorie hingegen wird der Volumenanstieg bei der Beurteilung der Mischungsintensität überhaupt nicht beachtet, weil die Kinder nicht über eine ausreichende Gedächtniskapazität verfügen, um neben der Intensität auch noch das Volumen zu berücksichtigen. Die Intensität der Mischung nimmt nur deshalb zu, weil sich die Farbteilchen aus den beiden Ursprungsgläsern addieren.

Kritisch muss hier allerdings angefügt werden, dass sich Jäger bei seinen Interpretationsversuchen stets nur auf die Ergebnisse jeweils einer Kombination von Farbintensitäten bezog. Den angeblich nicht gefundenen u-förmigen Verlauf erklärte er sich mit der direkten Darbietung der intensiven Grösse. Dabei liess er aber den annäherungsweise u-förmigen Verlauf bei der Kombination unterschiedlicher Farbintensitäten ausser acht. Da aber bei beiden Bedingungen die gleiche intensive Grösse und dieselbe Urteilsskala benutzt wurden, konnten die häufiger angewendeten Durchschnittsregeln bei der gleichen Intensitätskombination zumindest nicht ausschliesslich an der Aufgabenpräsentation gelegen haben. Auch die Argumentationen mit der Volumen- und Teilchentheorien greifen zu kurz, weil sie die widersprüchlichen Ergebnisse von gleicher und unterschiedlicher Farbintensitätskombination nicht zu erklären vermögen. In diesem Fall bezog sich Jäger auf das schlechtere Abschneiden bei der Mischung unterschiedlicher Farbintensitäten und ignorierte die deutlich besseren Resultate bei der Mischung von zwei gleich hellen Flüssigkeiten.

Experiment 2: Nachweis der Integrationsregeln bei der Kombination von zwei Flüssigkeiten mit unterschiedlichen Farbintensitäten

War bisher bereits von Durchschnitts- und Additionsregeln die Rede, so wurde dabei vorausgesetzt, dass die beiden Ursprungsintensitäten miteinander in Beziehung gesetzt werden. Dabei war aber noch keineswegs bewiesen, ob die Additionsregel beispielsweise überhaupt aus der additiven Verknüpfung der beiden Ursprungsintensitäten resultiert. Für diesen Nachweis bedurfte es einer systematischen Variation der einzelnen Ursprungsintensitäten, wie es Jäger in einem weiteren Experiment auch gemacht hat (Jäger, 1996; Jäger & Becker, 1997; Jäger & Wilkening, 2001). Das Vorgehen entsprach der Methode des funktionalen Messens im Rahmen der Informationsintegrationstheorie von Anderson (1982, 1996), wodurch eine Identifizierung und Unterscheidung verschiedener verwendeter Verknüpfungsregeln wie etwa die Addition, die Durchschnittsbildung oder die Multiplikation möglich wird. Weil die Methode aber eine systematische Variation der Ursprungsintensitäten voraussetzt, konnten nur die Mischungsregeln bei der Kombination unterschiedlicher Farbintensitäten – aber nicht bei der Kombinationen bei zwei gleichen Farbintensitäten – festgestellt werden.

Jäger untersuchte die Mischungsregeln bei der Kombination unterschiedlicher Farbintensitäten mit einem vollständig durchkombinierten 2x2-Versuchsplan. Die beiden Fak-

toren bestanden aus einer hohen und einer niedrigen Farbintensität, die jeweils in zwei verschiedenen hellen Ausprägungen dargeboten wurden. Als Urteilsskala diente diesmal eine 32-teilige Farbskala. Um die Anwendung der Volumenregel (siehe S. 12) auszuschliessen, mussten sich die Kinder vorstellen, dass die Mischungsflüssigkeit halbiert würde. Da die Lösungsstrategie nach der Volumenregel nicht zu einer Steigerung der Farbintensität führen kann, kann einem Mischungsurteil nach der Additionsregel nicht die Volumentheorie zu Grunde liegen.

Die Ergebnisse lieferten den Nachweis, dass die Kinder die beiden Faktoren (Ursprungsintensitäten) tatsächlich in ihr Mischungsurteil integriert hatten. Zudem konnten verschiedene Regeln bei der Integration der relevanten Grössen identifiziert werden, die sich unabhängig vom Alter deutlich unterschieden. Somit konnte gezeigt werden, dass der kognitive Prozess, der zum Mischungsurteil führte, bei allen Versuchspersonen unabhängig vom Alter gleich ablief. Damit war der Nachweis geliefert, dass von Durchschnittsbildung und Addition im Sinne von Informationsintegrationsregeln gesprochen werden darf.

Durch Regelzuweisung der Mischungsurteile wurde folgendes Entwicklungsmuster gezeichnet: Die jüngsten Kinder urteilten sehr häufig inkonsistent, die Acht- und Zehnjährigen urteilten am meisten gemäss der Additionsregel und bei den Zwölfjährigen dominierte deutlich die Durchschnittsbildung. Die Entwicklung der korrekten Durchschnittsbildung präsentierte sich im Entwicklungsverlauf stetig ansteigend. Die vielen inkonsistenten Urteile der jüngeren Kinder wurden im Hinblick auf das Urteilsverhalten, nachdem sie die ganze Urteilsskala nutzten, so interpretiert, dass sie der Erwartung, möglichst die ganze Skala zu nutzen, zu entsprechen versuchten. Würde man die Stufen der Urteilsskala reduzieren, so würden auch unter den Sechsjährigen mehr Addierer erwartet, wie sich dies in vielen Studien zur Addition von extensiven Grössen gezeigt hatte (Anderson & Cuneo, 1978a; Wilkening, Levin & Druyan, 1987).

Auf die weitere theoretische Bedeutung des Nachweises von Integrationsregeln wird im weiteren Verlauf dieses Kapitels (S. 19ff.) noch ausführlicher eingegangen.

Experiment 3: Einfluss des Volumens (extensive Grösse) auf die Farbintensität (intensive Grösse)

In einem dritten Experiment von Jäger (1996, 1997; Jäger & Wilkening, 2001) wurde der Einfluss des Volumens auf die Farbintensität bei der Voraussage der resultierenden Farbintensität überprüft. Folgende Hypothese lag dem Experiment zu Grunde: Wenn die additive Verknüpfung der Farbintensitäten dadurch zu Stande kam, dass die intensive Grösse wie eine extensive behandelt wurde, so müsste sich die Variation des Volumens der Ursprungsflüssigkeiten auf die Intensität des Mischungsergebnisses auswirken. Oder einfacher ausgedrückt: Wenn das Volumen der Ursprungsflüssigkeit bei gleich bleibender Farbintensität grösser ist, sollte auch die Farbintensität des Mischungsergebnisses ansteigen. Um den Einfluss des Volumens auf die Einschätzung der resultierenden Mischungsintensität überprüfen zu können, musste das Volumen variiert werden. Untersucht wurde dies durch die Variation der Farbintensitäten (zwei hohe und zwei niedrige zu mischende Intensitäten; allerdings nicht durchkombiniert) und die Variation der Volumina der Ausgangsflüssigkeiten ($\frac{1}{2}$, 1 und 2 dl). Die Stichprobe setzte sich aus acht- und zehnjährigen Kindern sowie Erwachsenen zusammen.

Da sowohl Kombinationen mit gleichen und als auch mit unterschiedlichen Volumina zu beurteilen waren, wurden diese bei der Auswertung getrennt behandelt.

Ergebnisse bei den Kombinationen mit gleichen Volumina

Bei der Aufgabe der Kombination gleicher Volumina wirkt sich das Volumen der Ausgangsflüssigkeiten normativ nicht auf die resultierende Farbintensität aus. Der Durchschnitt ergibt sich durch den arithmetischen Mittelwert der beiden Farbintensitäten. Dennoch wurde bei dieser Untersuchung beobachtet, dass die meisten Kinder – aber auch einige Erwachsene – die extensive Grösse bei ihrer Wirkung auf die intensive Grösse berücksichtigten. Je grösser die Volumina der Ursprungsflüssigkeiten, umso höher wurde die resultierende Farbintensität eingeschätzt, obwohl diese normativ immer gleich geblieben wäre. Die dominierende Vorstellung der Versuchspersonen bestand darin, dass die Farbintensität einer Mischung vom Volumen der Flüssigkeiten, aus denen sich die Mischung zusammensetzt, abhängig ist.

Nebst dem Faktor Volumen war allerdings auch die Farbintensität signifikant. Es zeigten sich folgende zwei Entwicklungstendenzen: Der Einfluss des Volumens nahm mit dem Alter ab, obwohl er auch bei den Erwachsenen nicht vollständig verschwand, und der Einfluss der Farbintensität wurde mit zunehmendem Alter grösser. Bei den Zehnjährigen war bereits eine Berücksichtigung der Farbintensitäten festzustellen.

Auffällig war zudem, dass die Additionsregel praktisch nicht angewendet wurde. Die Einzelanalysen ergaben, dass die Kinder v. a. inkonsistent urteilten und zu einem Drittel die Durchschnittsbildung anwendeten, wobei auch hier noch der Volumeneinfluss bemerkbar war. Diese Ergebnisse liessen den Schluss zu, dass die Additionsregel durch die Volumenregel „je grösser das Volumen umso intensiver die Farbmischung“ ersetzt wurde.

Ergebnisse bei den Kombinationen mit unterschiedlichen Volumina

Nebst den Kombinationen gleicher Volumina mussten noch Kombinationen von 1 oder 2 dl der einen Flüssigkeit und je $\frac{1}{2}$ dl der anderen Flüssigkeit beurteilt werden. Durch die Variation der Volumina beider Ursprungsflüssigkeiten war eine direkte Aussage über den Einfluss einer extensiven Grösse auf die intensive bei der Beurteilung der resultierenden Farbintensität möglich. Aus formaler Sicht wurde die Aufgabe schwieriger, weil sich die Farbintensität der Mischung aus dem gewogenen – und nicht mehr dem arithmetischen - Mittelwert der dargebotenen Farbintensitäten ergibt.

Bei den Ergebnissen zeigte sich ein fast identisches Bild wie bei den Kombinationen gleicher Volumina: Erneut wurde kaum addiert, bei den Kindern wurden vorwiegend inkonsistente Urteile beobachtet. Im Unterschied zu den Kombinationen gleicher Volumina zeigte sich bei den Durchschnittsbildern hier aber eine korrekte Integration der beiden Grössen und nicht mehr einen Anstieg mit zunehmendem Volumen.

Die verbesserten Einschätzungen könnten allerdings an einer spezifischen Eigenheit der Stimulusdarbietung gelegen haben: Die Farbintensität des Glases mit dem grösseren Volumen, die bei der Bildung des Mittelwerts mit einem höheren Gewicht einging, könnte auch auf der Wahrnehmungsebene der dominantere Reiz gewesen sein. Auch eine blossе Ausrichtung auf den salienteren Reiz könnte die Kinder zur richtigen Lösung der Aufgabe geführt haben.

Als besonders bedeutsames Ergebnis kann gewertet werden, dass in diesem Experiment die Additionsregel zu Gunsten der Regel „je voller die Ursprungsgläser desto intensiver die Mischung“ aufgegeben wurde. Möglicherweise reichte die Verarbeitungskapazität der Kinder nicht aus, um zusätzlich noch eine Additionsregel anzuwenden. Denkbar ist aber auch, dass durch die Berücksichtigung und den Einbezug des Volumens es offensichtlich wurde, dass ein Addieren der intensiven Grösse nicht die richtige Lösung sein kann.

2.1.2 Bisherige theoretische Konzeptionen zur Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen

Theorie der u-förmigen Entwicklung

Die theoretische Konzeption des u-förmigen dreiphasigen Verlaufs der kognitiven Entwicklung von Strauss (1982c) stützt sich auf eine breite Basis empirischer Untersuchungen mit intensiven Grössen (Stavy, 1981; Stavy et al., 1982; Strauss, 1982a; Strauss & Stavy, 1982; Strauss & Klein, 1985).

In der ersten Phase der vier- und fünfjährigen Kinder wurden die Aufgaben zur Mischung von intensiven Grössen mehrheitlich (qualitativ) korrekt gelöst. Die verbalen Begründungen der Kinder wiesen auf ein intuitives erfahrungsgestütztes Verständnis der Aufgabe. Die häufigsten Antworten in dieser Phase basierten auf einer Gleichheits- oder Identitätsregel, nach der die Intensität gleich geblieben war, weil nichts dazugegeben wurde oder weil die Flüssigkeiten auch nach dem Zusammenleeren immer noch die gleichen waren.

In einer zweiten Entwicklungsphase, dem kindlichen Alter von sechs bis zehn Jahren zugeordnet, hingegen traten solche Antworten kaum mehr auf. Statt dessen neigten die Kinder mehrheitlich dazu, die intensiven Grössen zu addieren, was zu der falschen Antwort führte, das Mischungsresultat sei süsser oder wärmer als die beiden Ausgangsflüssigkeiten. Strauss beobachtete, dass die intensiven Grössen des Mischungsresultates und der Vergleichsflüssigkeit in eine direkte funktionale Beziehung gebracht werden (*direct functioning reasoning*), im Sinne von „mehr Zucker = süsser“ bzw. „mehr Kerzen = wärmer“. Das zentrale Problem bestand darin, dass die intensive Quantität dabei

wie eine extensive Grösse behandelt wurde. Diese *extensivity bias* (nach Jäger & Wilkening, 2001), d. h. die Präferenz, die intensive Grösse als eine extensive zu behandeln, könnte ihren Ursprung in der reichen Alltagserfahrung mit extensiven Grössen haben (*general metric sense for extensive quantities* nach Anderson & Wilkening, 1991; weitere Ausführungen zur Hypothese der *general-purpose adding rule* im Verlaufe dieses Kapitels ab S. 21). Aufgrund der häufigen und vielfältigen Erfahrung mit extensiven Grössen im Alltag versuchen die Kinder, die physikalischen Aufgaben auf der Basis von vertrauten Situationen aus der physikalischen Welt zu lösen, welche eben vorwiegend extensive Grössen beinhalten. Ein anderer Grund für die Präferenz der Additionsregel in Aufgaben aus der intuitiven Physik könnte auch in der Art der Vermittlung dieser mathematischen Operation im Schulunterricht gesehen werden, bei der die Lehrpersonen die Addition sehr häufig mit Beispielen von extensiven Grössen illustrieren. Ergebnisse einer Studie von Strauss, Ankori, Orpaz und Stavy (1977) unterstützen diese Vermutung. Ihr Vergleich von Grundschulkindern mit gleichaltrigen Kindern, die nicht zur Schule gingen, zeigte, dass letztere bei Temperaturentscheidungen die intensive Grösse viel seltener addierten als die Schulkinder.

In der dritten Entwicklungsphase der elf- bis dreizehnjährigen Kinder dominierten zunehmend bidimensionale und proportionale Begründungen, die wieder zu richtigen Lösungen führten. In diesem Stadium gelang es den Kindern, die intensive von der extensiven Grösse in ihren Voraussagen von Mischungsergebnissen zu unterscheiden.

Auf einer abstrakteren und tieferen Ebene entspricht die dreiphasige Entwicklung nach der Vorstellung von Strauss (1982b) einem Wechsel von einem früheren repräsentationalen System (Phase 1) zu einem späteren andersartigen System (Phase 3). In der Übergangsphase (Phase 2) interagieren diese beiden Systeme, was meistens zu einem temporären Fehlverhalten führt. Das früh zum Ausdruck kommende System ist qualitativer, intuitiver Art und umfasst die Eigenschaften der physikalischen Materie. Das spätere System ist numerisch und repräsentiert die kognitiven Funktionen (Operationen und Regeln).

Kritik gegenüber dieser Theorie kam beispielsweise von Jäger (1996), dessen Ergebnisse der Kombination gleicher Farbintensitäten – auch bei den Studien der Gruppe um Strauss wurden gleiche hohe Intensitäten gemischt – keinen u-förmigen Entwicklungs-

verlauf aufwiesen. Jäger führte den Leistungsrückgang bei den älteren Kindern auf die quantitative, indirekte Aufgabendarbietung zurück, welche den Kindern nahegelegt habe, das Schulwissen über mathematische Operationen anzuwenden. Auch nach Ahl, Moore und Dixon (1992; weitere Ausführungen dazu weiter unten ab S. 23) wird durch eine digitale symbolische Darbietung der Stimuli das analytische Denken gefördert. Intuitives Denken hingegen werde durch die direkte Darbietung der relevanten Aufgabenmerkmale begünstigt und bewirke keinen Leistungseinbruch.

Ein weiterer Einwand zur Theorie der u-förmigen Entwicklung stammt von Jäger und Wilkening (2001) und ist auf das *direct functioning reasoning* gerichtet: Bei ihrer Untersuchung mussten sich die Kinder vorstellen, dass die Mischungsflüssigkeit nach dem Zusammenleeren wieder halbiert würde. Es wurde angenommen, dass die Kinder, würden sie die intensive Grösse wie eine extensive (Volumen) behandeln, die intensive Grösse nach dem Zusammenleeren ebenfalls hätten halbieren müssen. Ein Vorkommen der Additionsregel könnte somit nicht durch das *direct functioning reasoning* erklärt werden. Ihre Resultate zeigten aber keinen Effekt der Halbierung des Volumens der Mischungsflüssigkeit, da die Additionsregel v. a. bei den acht- und zehnjährigen Kindern die vorherrschende Mischungsregel war. Allerdings kann hier kritisch eingewendet werden, dass es nicht garantiert war, dass die Kinder den letzten Schritt der Aufgabe, die Halbierung, überhaupt mitvollzogen hatten. Es ist vorstellbar, dass die kognitiven Anforderungen für die jüngeren Kinder bei dieser Aufgabe zu hoch waren, um die Aufgaben korrekt zu lösen.

Entwicklung der Informationsintegration bei intensiven Grössen

Ein anderer Ansatz zur Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen kommt aus der Theorie der Informationsintegration von relevanten Grössen. Mit diesem Ansatz kann bestimmt werden, ob die relevanten Informationen überhaupt berücksichtigt werden und nach welcher Regel sie verknüpft bzw. integriert werden. Im Falle der Mischung von intensiven Grössen sollten die beiden intensiven Grössen unter Berücksichtigung der extensiven Grösse verarbeitet werden. Die beiden intensiven Grössen dürfen also nicht isoliert von der extensiven Grösse (Volumen) einfach addiert werden, da ihre Grösse vom Volumen abhängt.

Angesichts der Komplexität des Prozesses der Durchschnittsbildung scheint es zunächst nahe liegend, dass die Addition ein erster Schritt in die richtige Richtung sein könnte. Für die Anwendung der korrekten Verknüpfungsregel der Durchschnittsbildung ist aber noch ein weiterer kognitiver Verarbeitungsschritt nötig: Die beiden addierten Grössen müssen noch in Bezug zur extensiven Grösse gesetzt werden. Im Falle von zwei gleich hohen extensiven Grössen ist dies der arithmetische Mittelwert bzw. eine Halbierung der Mischungsflüssigkeit. Bei unterschiedlich ausgeprägten extensiven Quantitäten gehen die extensiven Grössen mit einem Gewichtungsfaktor in die Verknüpfungsregel ein. Aus entwicklungspsychologischer Sicht wäre der Nachweis einfacherer Verknüpfungsregeln bei den jüngeren Kindern und komplexerer Regeln bei den älteren Kindern eine geeignete Erklärung für die Entwicklung der kognitiven Durchschnittsbildung.

Die Methode der Wahl zur Identifizierung von Verknüpfungsregeln ist das funktionale Messen nach Anderson (z. B. 1980). Methodische Voraussetzungen, um die Integrationsfunktion messen zu können, sind die Verwendung eines vollständig durchkombinierten faktoriellen Untersuchungsdesigns und einer quantitativen Antwortskala.

Nachweis von verschiedenen Verknüpfungsregeln

Jäger (1996; Jäger und Wilkening, 2001; vgl. weiter oben Experiment 2, S. 13f) konnte mit der Methode des funktionalen Messens nachweisen, dass die beiden relevanten Informationen der intensiven Grösse tatsächlich miteinander verknüpft werden. Er gab den Kindern eine zweistufige hohe und niedrige Intensität vor, die er in einem vollständig durchkombinierten faktoriellen Versuchsdesign präsentierte. Es zeigte sich, dass sich unter den Versuchspersonen zwei Gruppen abhoben, welche die beiden abgestuften Intensitäten exakt berücksichtigten und diese auch jeweils konsistent entweder nach der Durchschnitts- oder Additionsregel verknüpften. Die dritte grosse Gruppe urteilte inkonsistent ohne erkennbare durchgängige Strategie. Während die Sechsjährigen nach Zuweisung zu Strategien vorwiegend inkonsistent urteilten, herrschte bei den Acht- und Zehnjährigen eine Dominanz der Additionsregel vor. Die Durchschnittsbildung wurde zwar mit zunehmendem Alter vermehrt angewendet, war aber erst bei den Zwölfjährigen deutlich vorherrschend gegenüber den anderen Regeln. Zumindest im Kontext einer spezifischen Aufgabenstellung konnte nachgewiesen werden, dass sich die Art der Verknüpfung in der Entwicklungsspanne von sechs bis zwölf Jahren verändert hatte und zwar in der Reihenfolge von Inkonsistenz über Addition zu Durchschnittsbildung.

Im Hinblick auf andere Studien aus dem Bereich der intuitiven Physik erstaunt es wenig, dass gerade die Additionsregel als Vorläufer der Durchschnittsregel beobachtet wurde, da dies auch in anderen Aufgabensituationen regelmässig der Fall war.

Hypothese einer allgemein verfügbaren Additionsregel (*general-purpose adding rule*)

Die Beobachtung, dass jüngere Kinder die einfachere Additionsregel anstelle einer scheinbar komplizierteren Regel anwenden, ist nicht neu. Es wurde sogar eine generelle starke Präferenz für Additionsregeln anstelle von normativen Regeln wie etwa der Multiplikationsregel beobachtet (Anderson & Cuneo, 1978a; Wilkening, 1979, 1981). Die starke Tendenz zur Verallgemeinerung der Additionsregel führte zur Formulierung der Hypothese einer allgemein verfügbaren Additionsregel bei jüngeren Kindern (*general-purpose adding rule*; Anderson & Cuneo, 1978a; Anderson, 1980). In dieser frühen Neigung zur Generalisierung des Additionsprinzips wird der Grund gesehen, warum Kinder intensive Grössen addieren, auch wenn sie höchst unangemessen sind und zu ungewöhnlichen Voraussagen führen. Nach der Informationsintegrationstheorie wird angenommen, dass die Kinder die additive Verknüpfungsregel deshalb früher beherrschen, weil sie mit extensiven Grössen, bei denen diese Regel zur Anwendung kommt, weit mehr Erfahrung gesammelt haben als mit intensiven Grössen. Bei der Beurteilung der Mischungsintensität wird die bereits verfügbare Additionsregel, die in anderen Zusammenhängen erworben wurde, angewendet.

Allerdings gibt es auch Befunde, die gegen die Hypothese einer allgemein verfügbaren Additionsregel sprechen oder zumindest einen allfälligen Anspruch auf Universalität einschränken: Sowohl die Anwendung der Durchschnitts- als auch der Multiplikationsregel wurde bei viel jüngeren Kindern in anderen Aufgabensituationen nachgewiesen (Multiplikation: Anderson & Wilkening, 1991; Wilkening & Anderson, 1991).

Durchschnittsregel bei jüngeren Kindern durch Aktivierung eines kognitiven Schemas

In einer Studie von Butzin und Anderson (1973) waren bei der Attraktivitätseinschätzung einer Kombination von verschiedenen attraktiven Spielzeugen sogar 5-Jährige zur Integration verschiedener Informationsteile mittels Durchschnittsbildung befähigt. Wei-

tere Studien im Bereich der sozialen Kognition konnten bei jüngeren Grundschulkindern ein Verstehen und Anwenden der Durchschnittsbildung nachweisen (z. B. Singh, Sidana & Saluja, 1978a, 1978b, Singh, Sidana & Srivastava, 1978).

Ein weiteres Argument, welches gegen eine generelle Unfähigkeit zur Durchschnittsbildung der jüngeren Grundschul Kinder spricht, stammt von Jäger (1996): Er konnte zeigen, dass fast ausschliesslich addierende Achtjährige die korrekte Mischungsregel bereits nach zweimaliger Rückmeldung durch die Präsentation des richtigen Resultates gelernt hatten (Jäger, 1996; Experiment 4).

Anderson (1981, 1987) schlug vor, dass das intuitive richtige Lösen mittels Durchschnittsbildung durch die Aktivierung eines kognitiven Schemas, dem mentalen Ausbalancieren, zu Stande kommt. Dieses mentale Ausbalancieren lässt sich gut anhand des Beispiels der Bestimmung des Schwerpunktes bei einer Balkenwaage illustrieren. Eine Aktivierung dieses Durchschnittsschemas in den Aufgabenstellungen der Studien im sozialen Bereich (s. o.) scheint sehr plausibel zu sein und könnte erklären, warum die Kinder bereits zur Durchschnittsbildung befähigt waren. Die Aktivierung eines kognitiven Schemas im Sinne des mentalen Ausbalancierens wird auch zur Erklärung der Befunde einer Untersuchung von Reed und Evans (1987) herangezogen. In ihrem Experiment zeigte sich, dass viele erwachsene Probanden eine Säuremischungsaufgabe erst dann lösen konnten, nachdem sie eine Aufgabe zur Temperaturmischung lösen mussten. Durch die Temperatureaufgabe, die aus einem weit vertrauteren Bereich stammt als die Säuremischungsaufgabe, konnte das kognitive Schema auf ein anderes Anwendungsgebiet transferiert werden. Offen blieb aber die Frage, ob sich das mentale Balancieren auch auf weniger ähnliche oder abstraktere Aufgabenstellungen wie etwa zum proportionalen Denken oder zur Bildung des gewogenen Mittelwertes übertragen lässt. So waren bei Jägers Experiment (1996) zur Lernfähigkeit durch Rückmeldung auch nur Aufgaben gestellt worden, die die Bildung des arithmetischen Mittelwertes erforderten. Um zu prüfen, ob das kognitive Schema des mentalen Balancierens auch in abstrakteren Aufgaben aktiviert wird, müsste eine Aufgabe gestellt werden, die nur mittels gewogenem Mittelwert gelöst werden kann (z. B. durch Variation des Volumens).

Diese Beobachtungen müssen nicht im Widerspruch zur Hypothese der allgemein verfügbaren Additionsregel stehen, wenn man die Schwierigkeiten der Kinder, die Durch-

schnittsbildung anzuwenden, als Problem des Zugriffs versteht. Auch Jäger und Wilkening (2001, p. 343) hatten vermutet, dass die Kinder Aufgaben aus dem sozialen Bereich von denjenigen aus dem physikalischen Bereich unterscheiden und sie auf verschiedenen Arten zu lösen versuchen. Die Verwendung der allgemein verfügbaren Additionsregel, die aufgrund der weit reichenden Erfahrung mit extensiven Grössen vertraut ist, beschränkt sich demnach zumindest auf physikalische Aufgaben, wie etwa das Lösen von Mischungsintensitätsaufgaben.

Das Vorkommen oder Ausbleiben der Durchschnittsregel ist nicht nur abhängig vom Anwendungsbereich, sondern es lassen sich auch Unterschiede im Denkvorgang der Durchschnittsbildung selbst vermuten. Das Schema des mentalen Ausbalancierens scheint beispielsweise nicht die geeignete Lösungsstrategie bei abstrakteren Problemen zu sein. Ahl, Moore und Dixon (1992) hatten zwischen intuitiven und analytischen Denkvorgängen unterschieden. Nach dieser Vorstellung entspricht die Durchschnittsbildung durch Aktivierung des Schemas des mentalen Ausbalancierens dem intuitiven Denken.

Durchschnittsbildung auf der Basis intuitiven vs. analytischen Denkens

Anderson selbst, der das kognitive Schema des mentalen Ausbalancierens vorgeschlagen hat, unterschied – in Anlehnung an die Darbietung der Mischungsaufgaben – zwischen analogem und digitalem Prozess der Durchschnittsbildung (Anderson, 1982). Bei Ahl, Moore und Dixon (1992) wurde der gleiche Sachverhalt mit der Dichotomie intuitives vs. analytisches Denken bezeichnet (vgl. Hammond, 1982; Kahneman & Tversky, 1982)¹.

Das intuitive Denken ist nach Ahl, Moore und Dixon (1992, p. 82) charakterisiert durch Urteilsbildungen, die auf nicht-numerischer Information basieren oder nicht durch explizite Kalkulation erreicht werden. Das intuitive Denken ist geprägt durch alltagsnahes Wissen und führt bei intuitiven Aufgaben zur richtigen Lösung. Analytisches Denken

¹ Dichotomisierungen von Denkart wurden schon viel früher vorgenommen: Piaget und Inhelder (1977) unterschieden zwischen qualitativen und quantitativen Konzepten. Die Terminologie von intuitivem und analytischem Denken geht auf Brunswik (1956) zurück, nach dessen Vorstellung sich auch das erwachsene Denken zwischen diesen beiden Polen eines kognitiven Kontinuums bewegt.

hingegen zeichnet sich durch die Anwendung von schulgelernten Regeln aus und kommt in symbolischer Aufgabengestaltung zum Einsatz. Als typisches Beispiel einer symbolischen Aufgabengestaltung können die Aufgaben von Strauss und Mitarbeitern gelten (z. B. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982).

Ahl, Moore und Dixon (1992) postulierten aufgrund der Beobachtung von Lösungsstrategien bei der Temperaturmischungsaufgabe zwei gleichzeitig ablaufende Entwicklungsprozesse: Der erste betrifft die wachsende Fähigkeit, sich der Regeln bewusst zu werden und auf andere Inhalte zu übertragen (im Sinne des *structure mapping* nach Gentner, 1983). Der zweite Entwicklungsprozess bezieht sich auf den zunehmend leichteren Wechsel zwischen intuitivem und analytischem Denken. Ahl, Moore und Dixon (1992) betrachten die Fähigkeit zum Wechsel als einen wichtigen Aspekt in der kognitiven Entwicklung. Dieser Wechsel vollziehe sich nur spontan, wenn die Gedächtnisverfügbarkeit der intuitiven Repräsentationen hoch sei. Befunde zur Schwierigkeit von Grundschulkindern, den Wechsel von verschiedenartigen Aufgaben spontan zu vollziehen, gibt es aus vielen entwicklungs- bzw. pädagogisch-psychologischen Studien (z. B. Stavy & Berkovitz, 1980; Strauss, 1981). Die Unfähigkeit zum spontanen Wechsel zwischen den beiden Denkart bei jüngeren Grundschulkindern könnte der Grund für den häufig gefundenen u-förmigen Verlauf der korrekten Lösungen von Mischungsaufgaben sein: Quantitative Darbietungen fördern das analytische Denken, das bei Schulkindern schon da ist, und führt durch die unangebrachte Anwendung zu einem Leistungseinbruch bei jüngeren Schulkindern. Vorschulkinder können aber noch nicht auf Schulwissen zurückgreifen und lösen deshalb die Aufgabe intuitiv richtig.

Ob eine Aufgabe analytisch oder intuitiv gelöst wird, hängt entscheidend von den Aufgabencharakteristika der Aufgabe selber ab. Hammond (1982; Hammond, Hamm, Grassia & Pearson, 1987) entwickelte hierzu eine Theorie über die Art, wie Aufgabencharakteristika das Ausmass beeinflussen, ob eine Person eine Aufgabe intuitiv oder analytisch lösen wird.

Entwicklung der Durchschnittsbildung als adaptiver Prozess

Auch beim Ansatz des adaptiven Denkens bzw. der *assemblage theory* von Anderson und Wilkening (1991; Wilkening, 1989) spielen die Aufgabencharakteristika eine zentrale Rolle beim Entscheid, welche Lösungsstrategie angewendet wird. Der Fokus dieses

Ansatzes richtet sich in erster Linie auf die z. T. sehr unterschiedlichen Leistungen der gleichen oder gleichaltrigen Kinder in Aufgaben mit z. T. nur geringer Aufgabenvariation. Kinder, die nicht alle Aufgaben auf der gleichen Stufe lösen können, demonstrieren dabei in erster Linie die Zielgerichtetheit und Flexibilität des kindlichen Denkens. Das Wissen von Kindern wird nicht mehr als „reines“ Konzept verstanden, sondern als eine Ansammlung (*assemblage*) von Merkmalen verschiedener Fähigkeiten, die in Abhängigkeit von Aufgabensituationen zusammenwirken.

Somit versteht sich diese Sichtweise klar als Gegentheorie zu den stufentheoretischen Entwicklungskonzeptionen, welche sog. horizontale Verschiebungen (*décalage horizontale*), die sich dadurch äussern, dass nicht alle Aufgaben auf der gleichen Stufe gelöst werden können, nicht erklären können. Die radikale Vorstellung eines Konzepts im Sinne der bereichsübergreifenden Stufentheorie von Piaget, das nach Erwerb immer stabil sei, gilt auch in anderen Bereichen, wie zum Beispiel bei der Begriffsbildung, als überholt. Diese Schlussfolgerungen decken sich auch mit anderen modernen Theorien zur Entwicklung von Konzepten, wie etwa die Theorie zur Konzeptbildung von Barsalou (1987, 1993).

Der Ansatz des adaptiven Denkens richtet sich aber auch klar gegen die neueren liberalisierten Stufentheorien. Solche Theorien räumen zwar eine horizontale Verschiebung beim Konzepterwerb ein, halten aber daran fest, dass es in einem jeweiligen Inhaltsbereich ein Konzept in Reinform gibt (z. B. Bruner, 1964; Bruner, Olver & Greenfield, 1971; Case, 1985; Kosslyn, 1978; Siegler, 1986). Das von den spezifischen störenden Aufgabenanforderungen losgelöste Konzept werde in der Entwicklungsabfolge nur anders repräsentiert, z. B. zunächst sensumotorisch, dann anschaulich und schliesslich logisch propositional. Im Zuge dieser Sichtweise gelangte die sog. Vereinfachungsstrategie zur häufigen Anwendung und grosser Popularität. Dabei galt es, spezifische Aufgabenanforderungen wie die Belastung des Kurzzeitgedächtnisses oder die mangelnde Vertrautheit mit dem Versuchsmaterial zu minimieren, um das „reine Konzept“ aufzuzeigen. Die Anwendung der Vereinfachungsstrategie führte dazu, dass die Altersgrenzen, die von Piaget gesetzt wurden, deutlich nach unten verschoben wurden (*early competence look*; z. B. Gelman & Baillargeon, 1983). Anderson und Wilkening (1991) äussern sich aber kritisch zu den Schlussfolgerungen einiger Vertreter der radikalen *early-*

competence view, dass junge und noch jüngere Kinder – wenn man nur die richtigen Aufgaben stellt – über die gleichen Konzepte wie ältere Kinder verfügen.

Krist und Wilkening (1991) wiesen auf diverse Untersuchungen, die zeigen konnten, dass das Wissen im gesamten Entwicklungsverlauf auf verschiedenen Ebenen repräsentiert ist und diese Ebenen oft zusammenwirken oder sich parallel entwickeln. Dadurch sei die Vorstellung von postulierten Entwicklungssprüngen zwischen qualitativ voneinander abgegrenzten Repräsentationsebenen nicht haltbar. Auch die Eingrenzung auf eine spezifische Inhaltsdomäne reiche nicht aus, da eine Person auch hier fast nie nur ein einziges *developmental level* habe (Fischer, 1991; nach Krist & Wilkening, 1991). Anderson und Wilkening (1991) gehen noch weiter und mutmassen, ob die reinen Konzepte, nach denen gesucht wird, unnatürlich sind und vielleicht gar nicht existieren.

Nach dem Ansatz des adaptiven Denkens kann man sich repräsentiertes Wissen eher so vorstellen, dass bestimmte Arten von Funktionswissen wie etwa die allgemein verfügbare Additionsregel oder der *general metric sense for extensive quantities* in bestimmten Aufgaben aktiviert werden. Das Funktionswissen wird adaptiv eingesetzt und dessen Aktivierung ist abhängig von spezifischen Aufgabensituationen. Wenn Kinder noch nicht über das notwendige physikalische Wissen verfügen, um eine Aufgabe zu lösen, so versuchen sie mithilfe ihres Vorwissens und den Hinweisreizen der Darbietung eine Lösungsregel zu finden. Die Präsentation einer Aufgabe kann die Prozesse, die bei der Informationsverarbeitung beteiligt sind, erheblich beeinflussen. Auch Aebli (1984; nach Wilkening, 1989) konnte in einem anderen Zusammenhang zeigen, dass die Charakteristika einer Aufgabe „ad hoc“ bestimmt werden und entscheiden, welche Wissensformen zum Lösungsversuch herangezogen werden. Aus dieser Sicht betrachtet wird das Aufdecken von reinen Konzepten bedeutungslos, interessanter ist vielmehr das adaptive Denken von Kindern, verstanden als Anpassung der Denkvorgänge an wechselnde situationale Einflüsse.

In diesem Sinne sind auch additive Lösungsstrategien keine rigiden Fehlvorstellungen bzw. kein Misskonzept mehr, sondern sie tauchen nur in Abhängigkeit von der Aufgabe auf. So verschwand beispielsweise die Additionsregel zu Gunsten einer Volumenregel bei einer Aufgabe mit Volumenvariation (Jäger & Wilkening, 2001; Exp. 2; siehe oben,

S. 15ff.). Dieses Experiment verdeutlicht den Nutzen der sog. Komplikationsstrategie in Ergänzung zur Vereinfachungsstrategie.

Theoretische und methodische Implikationen des Ansatzes des adaptiven Denkens bestehen darin, die spezifischen Aufgabenanforderungen nicht mehr als unerwünschte Störgrösse zu betrachten, sondern systematisch einzubeziehen. Das bedeutet z. B. die Aufgabenanforderungen zu variieren und die Aufgaben in verschiedene Kontexte zu setzen. Forschungsleitend soll die Frage sein, ob bestimmte Regeln robust sind oder ob sie abhängig von situationalen Bedingungen sind. Falls die kindlichen Strategien von der Aufgabendarbietung abhängig sind, sollen die situationalen bestimmenden Einflüsse identifiziert werden.

Folgende situationalen Einflüsse wurden bisher im Zusammenhang mit der kognitiven Durchschnittsbildung bei intensiven Grössen identifiziert:

1. Symbolische und numerische Aufgabendarbietung induziert den *general metric sense for extensive quantities*

Die frühen Forschungen von Strauss und Mitarbeitern zur kognitiven Durchschnittsbildung bei intensiven Grössen boten den Kindern die relevante Information stets in symbolischer und numerischer Weise dar. Die Temperatur wurde durch die Anzahl Kerzen oder die Süsse in Form von Anzahl Zuckerstücken oder Löffeln präsentiert. Stellt man sich eine solche Aufgabe vor, so wird verständlich, dass die intensive Quantität addiert wird, ohne dabei die gleichzeitige Volumenzunahme zu berücksichtigen. Jäger verwendete deshalb in diversen Studien die intensive Grösse in Form der Farbhelligkeit als direkt wahrnehmbare Grösse. Für Jäger (1996) war der beobachtete Rückgang der Additionsregel bei der Mischung von zwei gleich hellen Flüssigkeiten der Beweis dafür, dass die Anwendung der Additionsregel durch die numerische Darbietung der relevanten Grösse induziert wurde.

2. Wahlaufgaben provozieren schnelle und weniger reflektierte Urteile

In den Studien von Strauss und Mitarbeitern (z. B. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982) wurden aber nicht nur symbolische Darbietungen der intensiven Grösse sondern auch Wahlaufgaben (*forced choice-task format*) verwendet. Bei der

Wahlaufgabe müssen die Versuchsteilnehmer ihre Entscheidung aufgrund von drei Antwortmöglichkeiten treffen. Es erscheint plausibel, dass diese Vorgabe zu weniger reflektierten Urteilen führt, als wenn die Aufforderung besteht, eine sehr präzise Einschätzung zu machen. Aus diesem Grund hatten die Versuchsteilnehmer in der Studie von Jäger und Wilkening (2001) ihre Einschätzungen des resultierenden Mischungsresultates auf einer 32-teiligen Farbintensitätsskala zu machen. Die Ergebnisse ihres Experimentes zeigten aber kein Verschwinden der Additionsregel, da die Mehrheit der acht- bis zehnjährigen Kinder die intensiven Grössen addierte, anstatt nach der Durchschnittsbildung zu verknüpfen.

2.2 Theorie des *extinction paradigm* bzw. *magic experiment*

Das vorherige Kapitel hat die widersprüchlichen Befunde zur Entwicklung der kognitiven Durchschnittsbildung aufgezeigt und deutlich gemacht, dass weitere Forschung auf diesem Gebiet zum besseren Verständnis der Sachlage nötig ist. Der Ansatz des adaptiven Denkens legt nahe, einen anderen Zugang zum Prozess der kognitiven Durchschnittsbildung zu suchen, um mehr über das Wissen, das in verschiedenen Aufgabensituationen angezapft wird, herauszufinden. In der vorliegenden Untersuchung wurde ein Zugang über die kindlichen Reaktionen auf gezauberte und nicht-gezauberte Mischungsresultate gesucht. Dieser Zugang tauchte in der Forschungsgeschichte in vereinfachter Form unter den Termini *extinction paradigm* und *magic experiment* mehrfach auf. In diesem Kapitel wird die Geschichte dieser Methode nachgezeichnet und deren Möglichkeiten zur Erkenntnisgewinnung in der kognitiven Entwicklungspsychologie mit dieser Methode aufgezeigt.

2.2.1 Studien zum *extinction paradigm*

Die Methode, welche als *extinction paradigm* bezeichnet und v. a. in den sechziger und achtziger Jahren häufig angewendet wurde, bezieht sich vornehmlich auf die Untersuchung des kindlichen Verständnisses für Invarianz.

Das Verständnis für Invarianz basiert auf dem logischen Prinzip der Erhaltung (*conservation*), welches von Piaget formuliert wurde und die Fähigkeit bezeichnet, zu erkennen, dass eine bestimmte Menge quantitativ gleich bleibt, wenn das Aussehen wie z. B.

die Form verändert wird. Kinder, welche das Erhaltungsprinzip voll und ganz verstanden haben, wissen, dass sich eine Menge nicht ändern kann, solange nichts weggenommen oder hinzugefügt wird. Piaget entwickelte die klassische Erhaltungsaufgabe, um das kindliche Verständnis des Invarianzprinzips zu überprüfen. Hierzu wurden den Kindern zwei identische Mengen präsentiert, von denen eine z. B. durch Verformung verändert wurde. Das Kind hatte nun einzuschätzen, ob die Menge erhalten blieb oder sich verändert hatte. Das Setting der klassischen Erhaltungsaufgabe bestand demnach in einer Paarvergleichsmethode, welche aber vom Kind zusätzlich zum Verständnis für Invarianz auch eine transitive Schlussfolgerung abverlangte. Wie in verschiedenen Studien gezeigt wurde, sollten Kinder im Alter von fünf bis sechs Jahren über die Fähigkeit zur transitiven Schlussfolgerung in Erhaltungsaufgaben verfügen (Elkind & Schoenfeld, 1972; Übersichtsarbeiten von Field, 1987 und McEvoy & O'Moore, 1991; Goswami, 2001, S. 316ff.).

Mit der Methode des *extinction paradigm* wurde nun die Standarderhaltungsaufgabe so verändert, dass die Transformation manipuliert wurde, in dem z. B. unbemerkt Material weggenommen wurde. Nach der Transformation wurde demnach ein Stimulus präsentiert, der tatsächlich eine Veränderung der Menge zeigte. Es wurde davon ausgegangen, dass ein Kind mit vollständigem Verständnis des Erhaltungsprinzips sich dem unbemerkt manipulierten Stimulus widersetzen müsste. Stellte es aber aufgrund des eben gezeigten gegenteiligen Beweises sein eigenes Prinzip infrage, so sei das Verständnis für Erhaltung noch nicht vollständig ausgereift.

Die erste Studie, welche das *extinction paradigm* anwendete, geht auf Smedslund (1961) zurück. Dabei wurde überprüft, ob sich trainierte Kinder, welche die Einsicht in die Invarianz des Gewichtes bei Formveränderungen durch ein Training erlangten, und Kinder, welche von Anfang nach dem Erhaltungsprinzip urteilten, sich in der Aufgabenstellung nach dem *extinction paradigm* unterschieden. Dabei wurde ihre Reaktion auf die Präsentation einer offensichtlichen aber unbemerkt vollzogenen Verletzung der Gewichtskonstanz nach Verformung einer Knetmassenfigur (es wurde heimlich Knetmasse entfernt) beobachtet. Folgende zwei Reaktionen wurden unterschieden: Auslöschung (*extinction*) des Prinzips von Gewichtskonstanz („es ist leichter, weil es flacher wurde“) und Widerstand gegenüber der Präsentation unter Beibehaltung des Erhaltungsprinzips („es ist leichter, weil Knetmasse entfernt wurde“). Es zeigte sich, dass

sich die beiden Gruppen der Kinder, welche einerseits natürlicherweise und andererseits durch Training über das Verständnis von Gewichtserhaltung verfügten, in ihrer Reaktion unterschieden: Erstere reagierten mehr mit Widerstand gegenüber der manipulierten Präsentation einer Durchbrechung des Erhaltungsgesetzes als die trainierten Kinder. Allerdings war dieser Widerstand auch bei den Kindern mit natürlich erworbenem Erhaltungsverständnis nur bei einer Minderheit zu beobachten. Diese Auslöschung des Gesetzes bei denjenigen Kindern, welche in den Standarderhaltungsaufgaben bereits von Anfang an über das Erhaltungsprinzip verfügten, wurde in der Folge zum zentralen Thema der Studien mit dem *extinction paradigm*.

Scott A. Miller gibt in seinem Übersichtsartikel von 1976 (eine revidierte und ergänzte Ausgabe seines Review-Artikels von 1971) einen Überblick über verschiedene Studien aus den 60er und 70er Jahre zum *extinction of conservation*-Paradigma. Das zentrale Ziel von Untersuchungen mit dieser Methode bestand in der Überprüfung der Stabilität der kindlichen Erhaltungsprinzipien. Dabei wurde die Erhaltung der quantitativen Menge durch unbemerkt manipulierte Präsentationen absichtlich verletzt, um das bereits in klassischen Erhaltungsaufgaben gezeigte Verständnis für Invarianz herauszufordern. Um die Manipulationen möglichst überzeugend vorzuführen und damit das noch junge Verständnis vom regelhaften Funktionieren der Natur zu erschüttern, wurden manchmal Techniken von Zauberkünstlern übernommen.

Miller (1976) teilte die kindlichen Reaktionen auf unbemerkt präsentierte Verletzungen des Erhaltungsprinzips (v. a. Gewicht und Menge) sämtlicher verwerteter Studien in die folgenden zwei Kategorien ein:

- **Auslöschung (*extinction*):** Das durch unbemerkte Manipulation entstandene Ereignis wird als gültig betrachtet und das Erhaltungsprinzip tritt ausser Kraft. Die Auslöschung des Gesetzes durch Konfrontation mit einer offensichtlichen Nicht-Erhaltung eines bestimmten Gesetzes wurde als natürliches Rückkehren auf eine frühere Funktionsstufe interpretiert.
- **Widerstand (*resistance*):** Die Gültigkeit des nicht erhaltenden Ereignisses wird dementiert und das Gesetz bleibt erhalten. Wird bei der Konfrontation mit einer offensichtlichen Durchbrechung des Erhaltungsgesetzes mit Widerstand gegen-

über dem Ereignis reagiert, so wird postuliert, dass das Verständnis des Erhaltungsprinzips voll und ganz ausgereift ist.

Miller (1976) betonte aber die Mängel bezüglich Transparenz und Vergleichbarkeit der verbalen Erklärungen der Kinder zu den dargebotenen Ereignissen, welche dieser Kategorisierung zu Grunde lagen. Bei der Systematisierung dieser kindlichen Erklärungen war zudem zu beobachten, dass die Kinder generell Schwierigkeiten hatten, logische Erklärungen zur Entstehung der Ereignisse zu geben.

Die theoretische Erwartung des Widerstandes bei Kindern, die in klassischen Erhaltungsaufgaben ein Verständnis für Invarianz zeigten, war auf der Basis von Piagets Stufenmodell breit abgestützt: Gemäss des postulierten grundlegenden und unidirektionalen Strukturwandels vom voroperatorischen zum konkret-operatorischen Stadium sowie des Entwicklungsgesetzes mit notwendigen Bedingungen (Piaget, 1971; Piaget & Szeminska, 1975/1941) war eigentlich nur ein Standhalten gegenüber dem gegenteiligen Beweis zu erwarten.

Miller (1976) kommt allerdings nach dem Vergleich von zwanzig verschiedenen Studien mit dem *extinction paradigm* zum Schluss, dass sich der Widerstand nur sehr selten zeigte. Zusammenfassend zeigten die Studien, dass zwischen 35-50% der kindlichen Reaktionen der Kategorie „Widerstand“ zugewiesen wurden. Miller äusserte zudem die Vermutung, dass der Widerstand durch mehrmalige Präsentation des manipulierten Ereignisses ansteigt. Als mögliche Gründe dafür, dass der erwartete Widerstand nicht im postulierten Ausmass gefunden wurde, nennt Miller (1976) die Unterschätzung des Widerstandes aufgrund diagnostischer Messfehler sowie die Infragestellung, ob die Auslöschung des Prinzips (*extinction*) in diesen Aufgaben einer tatsächlichen Rückkehr auf eine vorherige – im piagetschen Sinne auf die präoperatorische – Stufe gleichkommt. Nichtsdestotrotz können die gezeigten Unterschiede im Widerstand bzw. der Stabilität mit der Methode des *extinction paradigm* einzig so interpretiert werden, dass bestimmte Kinder tatsächlich ein solideres Verständnis des Erhaltungsprinzips haben als andere. Schon in den siebziger Jahren wurde vermutet, dass sich der Prozess des Erwerbs eines Konzeptes vom ersten rudimentären Verständnis bis zur vollständigen funktionsfähigen Reife möglicherweise über Jahre erstreckt (Flavell, 1971).

Trotz vielen fehlerbehafteten *extinction*-Studien hielt Miller an der Methode des *extinction paradigm* fest und sah sie damals als eine der wenigen verfügbaren Techniken, welche die Natur des „tiefen“ Verständnisses eines Konzepts (*in-the-system knowledge*), das sich über die Kindheit entwickelt, erfassen konnte und möglicherweise verschiedene Arten, wie man ein Konzept „haben“ kann, identifizieren konnte. Konkret sah Miller den Nutzen des *extinction paradigm* insbesondere in der Erfassung von Entwicklungs- und Konzeptunterschieden. Entwicklungsunterschiede könnten beispielsweise zwischen Kindern mit natürlichem Erhaltungsprinzip und Kindern mit eben erlerntem Verständnis für Invarianz (vgl. Smedslund, 1961) oder zwischen Kindern, die das Prinzip eben erworben haben und älteren Kindern untersucht werden. Hinter beiden Vorschlägen steckt die Hypothese, dass Kinder mit einem stabileren bzw. reiferen Verständnis für Invarianz resistenter gegenüber scheinbaren Durchbrechungen des Erhaltungsprinzips reagieren. Geht man davon aus, dass der Erwerb eines Konzeptes einen längeren Prozess beinhaltet, so könnte mit dem *extinction paradigm* der Verlauf der Stabilität über verschiedene Entwicklungsstufen hinweg nachgezeichnet werden. Untersuchungen zu Konzeptunterschieden mit dem *extinction paradigm* könnten etwa von der Hypothese ausgehen, dass zu unterschiedlichen Zeitpunkten erworbene Konzepte unterschiedlich stabil sind oder dass zwei verschiedene Konzepte sich generell in ihrer Stabilität bei der Konfrontation mit gegenteiligem Beweis unterscheiden.

Nach kritischer Beleuchtung von zwanzig Studien mit der Methode des *extinction paradigm* schlug Miller (1976) folgende Anweisungen für maximal effektive Studien mit dieser Methode vor:

- Die Resultate der manipulierten Transformation sollten mit der typischen Überzeugung einer Vorstufe in der kognitiven Entwicklung übereinstimmen. Im Falle des Beispiels mit der Gewichtsveränderung eines Knetballs nach Verformung sollte das Gewicht leichter und nicht schwerer sein, wenn der Ball flacher ist als vorher.
- Die Erfassung der Erklärungen zum Ereignis sollte unmittelbar nach der Transformation erhoben werden, wenn das Ereignis noch im Blickfeld des Kindes ist.
- Die Manipulation sollte einen Trick beinhalten, der dem Kind nicht bekannt ist und folglich nicht erkannt werden kann. Die Präsentation des einem Naturgesetz widersprechenden Ereignisses sollte so echt wie möglich wirken.

2.2.2 Ausweitung des Anwendungsbereichs: *Magic experiment*

Aus der Zeit nach Mitte der siebziger Jahre sind keine Studien mit der Methode des *extinction paradigm* mehr auszumachen. Etwa in derselben Zeit des *extinction paradigm* wurden auch Untersuchungen mit ähnlichen Verfahren wie dem *magic paradigm* und dem *magic experiment* durchgeführt. Obwohl sich die beiden Methoden hinsichtlich der Manipulation des transformierten Ereignisses deutlich unterschieden, war ihnen eine erweiterte bzw. veränderte Erfassung der Reaktionen auf das präsentierte Ereignis im Gegensatz zur Methode des *extinction paradigm* gemein: Statt die verbalen Erklärungen der Kinder aufzunehmen und zu kategorisieren wurde die Überraschungsreaktion erfasst. Zudem wurden mit beiden Verfahren auch jüngere Kinder untersucht, die in den klassischen Invarianzaufgaben noch kein Verständnis für das Erhaltungsprinzip gezeigt hatten.

Mit der Methode des *magic paradigm* wurden mehrere Studien zum Verständnis von Invarianz bei Zahlen durchgeführt (R. Gelman, 1972; Marschark, Greenberg & Clark, 1983; Silverman, Rose & Phillis, 1979). Bei diesen Untersuchungen wurde die klassische Zahlenerhaltungsaufgabe nicht mehr durch ein offenes, sondern ein unbemerktes und unerwartetes Verändern der räumlichen Anordnung einer Reihe von Elementen (Auseinanderziehen oder Zusammenstossen) verändert. Hierbei zeigte sich, dass die Kinder früher als in der klassischen Versuchsanordnung ein Verständnis für Zahlenerhaltung zeigten. Wesentlich für die begriffliche Unterscheidung ist aber, dass beim *magic paradigm* im Unterschied zum *extinction paradigm* und dem *magic experiment* keine Manipulation im Sinne einer Verletzung des Naturgesetzes vorgenommen wurde.

Die Methode, welche *magic experiment* genannt wurde und erstmals von Mermelstein und Shulman (1967) angewendet wurde, beinhaltete eine Veränderung der klassischen Invarianzaufgabe beim Umgiessen von Flüssigkeiten durch heimliche Zugabe oder Entnahme von Flüssigkeit sowie eine veränderte Erfassung der Reaktion durch die Überraschungsreaktion nach Beobachtung der Transformation. Lindsay und Creedon (1985) verbesserten diese Methode durch Reduktion der Flüssigkeitszugabe auf ein realistischeres Volumen und einer verfeinerten Kategorisierung der Überraschungsreaktion. In dieser Untersuchung wurden erstmals neben physikalisch unkorrekten Transformationsereignissen, welche den kognitiven Operationen Zentrierung und Addition entsprachen,

auch physikalisch korrekte Resultate präsentiert. Konkret wurde den Kindern nach dem Umgiessen von Wasser von einem schmalen in ein breites Glas ein höherer, gleich hoher und niedrigerer resultierender Wasserstand präsentiert. Durch Zugabe von Wasser durch eine versteckte Verbindung zu einem zusätzlichen Behälter ins breitere Glas konnte das Kind den Transformationsvorgang beobachten, ohne die Manipulation sehen zu können. Lindsay und Creedons Hypothesen (1985) standen im Zeichen der piaget-schen Vorlage zum Invarianzverständnis bei Flüssigkeiten: Kinder im Kindergartenalter erwarten einen höheren Wasserstand und sind beim Eintreffen der Erwartung wenig überrascht, bei den beiden anderen Präsentationen hingegen sind sie gleichermassen überrascht. Die älteren Kinder aus der dritten Klasse erwarten im Sinne eines ausgereiften Invarianzverständnisses einen tieferen Wasserstand und reagieren dementsprechend nur bei den physikalisch unkorrekten Umschüttergebnissen mit Überraschung. Interessanterweise waren aber auch die jüngeren Kinder beim höheren Wasserlevel mehr überrascht als beim tieferen Level. Dies, obwohl dieselben Kinder in weiteren Aufgaben nicht immer im Sinne des Erhaltungsprinzips urteilten. Die Autoren vermuteten, dass die jüngeren Kinder möglicherweise bereits das Prinzip der Kovarianz zwischen den Dimensionen erkennen, ohne aber das Erhaltungsprinzip zu verstehen. Bei den Drittklässlern, welche in allen Erhaltungsaufgaben ein Verständnis für Invarianz zeigten, wurde beim korrekten Wasserstand keine Überraschung registriert. Ihnen wurde ein vollständig ausgereiftes Verständnis des Erhaltungsprinzips attestiert. Allerdings insistierten nach Beobachtung des Umgiessens mit gleich gebliebenem Wasserstand längst nicht alle auf dem Erhaltungsgesetz und zeigten auch kaum eine Überraschungsreaktion. Bei der Bedingung mit höherem resultierendem Wasserstand hingegen wurde häufiger mit Überraschung reagiert als beim gleich gebliebenen.

Der teilweise ausgebliebene Widerstand beim gleich gebliebenem Wasserstand, der einer Zentrierung – der postulierten Frühform des Invarianzverständnisses - entsprach, führte erneut zur Frage, ob das *extinction paradigm* dazu verleite, auf eine frühere (vor-operatorische) Stufe zurückzukehren. Es war allerdings auch denkbar, dass die Kinder in der Aufgabe mit dem *extinction paradigm* am Verständnis für reziproke Kompensation gescheitert waren, welche in den Standarderhaltungsaufgaben zum Verständnis des Erhaltungsgesetzes nicht zwingend verstanden werden müssen. Ausserdem wurde die postulierte zwingende logische Konsequenz bei vorhandenem Verständnis des Erhaltungsprinzips infrage gestellt (vgl. Hall & Kaye, 1978). Stattdessen wurde eher vermu-

tet, dass die logische Konsequenz des Erhaltungsprinzips erst in einem längeren Entwicklungsprozess erworben wird und das Verständnis des Erhaltungsprinzips einen längeren Stabilisierungsprozess durchläuft (vgl. Miller, 1973).

Lindsay und Creedon (1985) versuchten im Anschluss an ihre Untersuchung den Prozess des Erwerbs des Erhaltungsgesetzes in Teilstrategien zu zergliedern. Sie vermuteten bei den Kindern im Kindergartenalter, welche trotz dem Scheitern in Erhaltungsaufgaben mit Überraschung auf manipulierte Wasserstände reagierten, ein rudimentäres Verständnis für das Prinzip der Kovarianz der beiden Dimensionen Breite des Glases und Wasserstand. Bei den älteren Kindern der dritten Klasse, welche zwar in Standardaufgaben nach dem Erhaltungsprinzip urteilten, aber in der *extinction*-Bedingung auf die physikalisch unkorrekten Resultate nicht mit Überraschung reagierten, wird vermutet, dass ihnen das Verständnis der Erhaltung in der schwierigeren Rekonstruktionsbedingung noch fehlt. Allerdings äusserten sich die Autoren selbst kritisch zur Erfassung der Überraschungsreaktion aufgrund von konfundierenden Faktoren, die den Ausdruck von Überraschung unabhängig vom präsentierten Stimulus beeinflusst haben konnten.

Mit ihrer Studie konnten Lindsay und Creedon (1985) zeigen, dass die zusätzliche Präsentation des physikalisch korrekten Resultates sowie der Einsatz der Methode des *extinction paradigm* auch bei jüngeren Kindern sinnvoll waren. Es kann nicht nur die Reaktion auf das physikalisch unmögliche Resultat erhoben und interpretiert werden, sondern auch diejenige auf das erwartete Ereignis, die theoretisch genauso begründet ist wie die Reaktion auf ein nicht erwartetes Ereignis. Erst wenn in beiden Bedingungen konsistent reagiert wird, kann von einem vollständig ausgereiften Verständnis des Prinzips gesprochen werden. Zudem konnte bei jüngeren Kindern auf ein gewisses Verständnis des Erhaltungsprinzips geschlossen werden, das in der klassischen Erhaltungsaufgabe nicht gezeigt werden konnte.

2.2.3 Kritik an der Methode des *extinction paradigm* bzw. *magic experiment*

Scott A. Miller (1971, 1973, p. 48) hat bereits Anfang der siebziger Jahre auf mögliche Versuchsleitereffekte bei der Methode des *extinction paradigm* hingewiesen. Er kritisierte, dass die Möglichkeit von sozial gefügiger „Pseudoextinktion“ in frühen Studien zu gross war, und hoffte, diesen Druck auf die Versuchspersonen in seinen eigenen Stu-

dien durch verschiedene abhängige Variablen wie etwa die Erfassung der Überraschungsreaktion zu minimieren.

Genauer untersucht wurde die Rolle des Versuchsleiters in Studien nach der Methode des *extinction paradigm* bzw. *magic experiment* erst in den neunziger Jahren von Chandler und Lalonde (1994). Auch sie kritisierten insbesondere den Einfluss des Versuchsleiters, der durch seine Autorität die Kinder in ihrem noch jungen Verständnis bestimmter Naturgesetze oder Erhaltungsprinzipien viel zu stark verunsichere und zur vorschnellen Verwerfung ihrer Gesetze führe. Die Kinder seien dabei gar nicht in der Lage, zwischen ihren Erhaltungsgesetzen und den anscheinend diesen Gesetzen widersprechenden Darbietungen entschieden zu müssen, sondern sie würden sich in einer zwischenmenschlich viel delikateren Problemsituation befinden. Es bestehe ein Dilemma zwischen ihren eigenen Erhaltungsvorstellungen und ihren Vorstellungen über die Ehrlichkeit und Wahrhaftigkeit der Erwachsenen, welche diese Darbietungen demonstrieren. Sie sahen in der unkontrollierten Rolle des Versuchsleiters auch den Grund für die widersprüchlichen Resultate diverser Studien aus den früheren Jahren (vgl. oben S. 28ff.).

Um den Effekt des Versuchsleiters auf die Kinder zu überprüfen, wurde in der Untersuchung von Chandler und Lalonde (1994) die angenommene Glaubwürdigkeit des Experimentators, der die Durchbrechungen der Naturgesetze präsentierte, systematisch variiert. Hierzu nahmen die Versuchsleiter möglichst authentisch drei verschiedene Rollen ein: als Priester, als Psychologe (was der Versuchsleiter in psychologischen Experimenten ja meist ist) und als Zauberer. Dem Priester und dem Psychologen wurde hypothetisch eine höhere und dem Zauberer eine niedrigere Glaubwürdigkeit zugeschrieben. Untersucht wurden 9- bis 13-jährige Kinder, die in einem Vortest ein vollständiges Verständnis der Gewichtserhaltung gezeigt hatten. Wie schon bei Smedslund (1961) wurde auch hier das Verständnis der Gewichtserhaltung anhand eines Knetballs überprüft, dessen Gewicht sich nach Verformung scheinbar – durch Manipulation der Waage mit einem Magneten – verändert hatte. Nach der Beobachtung der Versuchsdurchgänge wurden die Kinder jeweils von einem anderen Experimentator nochmals zu ihren Erhaltungsgesetzen befragt. Um den längerfristigen Einfluss der Präsentation manipulierter Ereignisse zu überprüfen, wurde das kindliche Verständnis der Gewichtserhaltung nach zehn Tagen erneut erfragt. Die Autoren erwarteten folgende Reaktion beim Zauberer:

Verteidigung des Erhaltungsprinzips entweder durch Anzweifeln der Legitimität der Präsentation durch Annahme eines dahinter stehenden Tricks oder dadurch, dass sie sich als Zeugen von echter Magie oder Wunder sahen. Die Präsentation beim Magier sollte also im Post-Test und v. a. im verzögerten Post-Test (10 Tage später) keine bleibenden Zweifel an ihren Erhaltungsgesetzen hinterlassen haben. Beim Priester und Psychologen war das nicht so klar, da diese möglicherweise so vertrauensvoll erschienen, dass die ursprünglichen Gesetze infrage gestellt wurden. Es zeigte sich in der Untersuchung von Chandler und Lalonde (1994), dass die unmittelbaren Reaktionen auf die Präsentation einer Verletzung des Erhaltungsgesetzes Überraschung und Zweifel waren. Nach der zweiten Beobachtung übernahmen viele und beim dritten Mal alle die vorübergehende Aufhebung des Gesetzes und sagten das Gegenteil voraus. Dass alle Kinder nach der dritten Beobachtung ein physikalisch unkorrektes Ergebnis erwarteten, erstaunt wenig: Sie hatten ja eben gerade gesehen, dass sich das Gewicht durch Verformung des Versuchsleiters immer verändert hatte.

Erst die Antworten der Kinder im Post-Test nach zehn Tagen konnten im Hinblick auf die Hypothese ausgewertet werden: Es zeigte sich, dass die Kinder aus der Zauberer-Bedingung bedeutend häufiger wieder ein physikalisch korrektes Resultat voraussagten, als die Kinder der Priester- und Psychologen-Bedingung. Die Erhaltungsgesetze der Kinder, welche den Zauberer beobachtet hatten, wurden weniger bedroht als durch die Beobachtung eines Priesters oder eines Psychologen. Dem Zauberer wurde offenbar bereits im Alter von neun bis dreizehn Jahren zugeschrieben, dass er unter bestimmten Umständen die Gültigkeit der Naturgesetze (scheinbar) aufzuheben vermochte.

Angesichts der berechtigten Kritik von Chandler und Lalonde (1994) sowie der Schwierigkeit, die Reaktion der Kinder auf Verletzungen der Erhaltungsgesetze mit der Überraschungsreaktion erfassen zu wollen, erstaunt es nicht besonders, dass das *extinction paradigm* und das *magic experiment* in der neueren Forschung nicht mehr verwendet wurde. Abgesehen von der Schwierigkeit des Versuchsleitereffektes sowie der Wahl einer geeigneten abhängigen Variablen scheint die Idee, die Stabilität einer kognitiven Integrationsregel oder eines kognitiven Konzeptes mit der Reaktion auf ein gegenteiliges Ereignis zu erfassen, aber nach wie vor fruchtbar zu sein.

Die Methode, mit gezauberten bzw. getricksten Ereignissen den Zugang zum kindlichen Denken zu finden, wurde denn auch in anderer Form in den letzten zwei Jahrzehnten ausgiebig und erfolgreich angewendet. Die Reaktion auf ein unerwartetes Ereignis mit der Überraschungsreaktion zu erfassen, ist für Untersuchungen mit Säuglingen und Kleinkindern eine geeignete Methode den Zugang zum Denken vorsprachlicher Kinder zu finden. Besondere Berühmtheit und Bedeutung haben die vielen Studien mit der Methode der Erwartungsverletzung (*violation of expectation*) von Renée Baillargeon (z. B. Baillargeon, 1986, 1993, 1995) erlangt. Die Überraschung bzw. Irritation auf ein unerwartetes Ereignis wurde jeweils durch Messung von Blickzeiten erfasst: je länger der Säugling auf das Ereignis schaut, umso irritierter ist das Kind (Dishabituation). Voraussetzung, dass der Säugling überhaupt irritiert sein kann, müssen frühe Kausalitätserfahrungen sein, die zumindest zu einer rudimentären Erwartungshaltung führen. Mit dieser Methode demonstrierte Baillargeon beispielsweise die Erwartung von Objektpermanenz bei erst viereinhalb Monate alten Babys (1987).

Wie dieser Forschungszweig – wie auch schon die Studie von Lindsay und Creedon (1985) – deutlich gemacht hat, kann die Methode der Erfassung der Reaktionen auf physikalisch mögliche und unmögliche Ereignisse auch zur Entdeckung von Prinzipien oder Regeln bei jüngeren Kindern führen. Kritiker der Studien mit Vorschul- und Schulkindern haben aber klar gemacht, dass eine Veränderung der Rolle des Versuchsleiters in Erwägung gezogen und eine neue abhängige Variable zur Erfassung der Reaktionen gefunden werden muss.

Die Ergebnisse der Studie von Chandler und Lalonde (1994) sprechen dafür, in Experimenten mit dem *extinction paradigm* oder einem *magic experiment* den Versuchsleiter als Zauberer zu präsentieren. Die Verletzungen des Erhaltungsgesetzes sind bei der Darbietung durch einen Zauberer weniger bedrohlich für das Kind, da ja der Zauberer eben gerade eine Figur repräsentiert, die eigentlich als einzige zu (scheinbaren) temporären Durchbrechungen von Naturgesetzen befähigt ist. Die Figur des Zauberers erscheint ideal zur Plausibilisierung manipulierter Ereignisse, ohne das noch junge Verständnis der Kinder zu stark zu bedrohen. Da der Zauberer mit der scheinbaren Verletzung von Naturgesetzen spielen kann, scheint es besonders geeignet, dies auch als abhängige Variable zu benutzen. In der vorliegenden Untersuchung mussten deshalb die Versuchspersonen gezauberte und nicht-gezauberte Ereignisse danach einschätzen, in

welchem Ausmass sie gezaubert sind. Weicht das präsentierte Ereignis von der eigenen Erwartung ab, so ist es gezaubert. Entspricht das präsentierte Ereignis aber der eigenen Vorstellung vom normalen Funktionieren der Welt, dann ist es nicht gezaubert. Dies setzt allerdings voraus, dass auch die jüngeren Kinder – in der vorliegenden Untersuchung waren die jüngsten Kinder im Kindergartenalter – ein inhaltliches Verständnis vom Zaubern haben und verschiedenartige dargebotene Ereignisse den Kategorien „gezaubert“ und „nicht-gezaubert“ zuordnen können. Dass diese Voraussetzung gegeben ist, wird im nächsten Kapitel anhand von mehreren Studien belegt.

2.3 Kindliches Verständnis von gezauberten Ereignissen

In der vorliegenden Untersuchung mussten physikalisch mögliche und unmögliche Ereignisse u. a. auf einem Kontinuum zwischen „nicht-gezaubert“ und „maximal gezaubert“ eingeschätzt werden. Dies setzt voraus, dass auch sechsjährige Kinder wissen müssen, was „Zaubern“ heisst und sie in der Lage sein sollten, gezauberte Ereignisse von nicht-gezauberten zu unterscheiden. In diesem Kapitel wird zunächst der Frage nach der Unterscheidungsfähigkeit von physikalisch möglichen und unmöglichen Ereignissen von Kindern und deren Operationalisierungsmöglichkeit nachgegangen. Im weiteren Verlauf wird das kindliche Verständnis der inhaltlichen Kategorie „Zaubern“ beleuchtet. Wie sich gezeigt hat, eignen sich gezauberte Ereignisse auch zur Erfassung des kindlichen Kausalitätsverständnisses. So sind bei Kindern z. B. verschiedene Arten von magischen kausalen Mechanismen vorzufinden. Eine Möglichkeit zur Erfassung der kindlichen Kausalitätskategorien besteht im Festhalten verbaler Erklärungsversuche der Kinder zur Entstehung gezauberter Ereignisse. Allerdings ist diese Methode stark von den Verbalisierungsfähigkeiten der Kinder abhängig, weshalb für eine indirekte Erfassung des Kausalitätsverständnisses der Kinder über Machbarkeitszuschreibungen verschiedener Akteure plädiert wird.

2.3.1 Unterscheidungsfähigkeit zwischen gezauberten und nicht-gezauberten Ereignissen bei Vorschulkindern

Gezauberte und nicht-gezauberte Ereignisse werden in der Forschung – wie im vorherigen Kapitel bereits berichtet wurde – schon seit mehreren Jahrzehnten für den Zugang

zur kindlichen Unterscheidungsfähigkeit von physikalisch möglichen und unmöglichen Ereignissen eingesetzt. Mit dieser Methode sollte jeweils untersucht werden, in welchem Alter die Kinder erstmals ein bestimmtes getrickstes Ereignis für unmöglich hielten. Waren die Kinder in der Lage, die unmöglichen Ereignisse von den möglichen zu unterscheiden, so wurde auf das Vorhandensein eines zumindest rudimentären Verständnisses des manipulierten physikalischen Prinzips geschlossen. In Abhängigkeit vom Alter der Kinder variierte das untersuchte Naturgesetz sowie die Art und Weise, wie die Reaktionen bzw. die Antworten der Kinder erfasst wurden.

Wie bereits erwähnt, wurde bei Säuglingen der nonverbale Zugang zum Verständnis von physikalischen Prinzipien – wie etwa das Verständnis, dass Materie nicht durchdrungen werden kann – über die „Überraschungs-/Irritationsreaktionen“ auf Präsentationen möglicher und unmöglicher Ereignisse gesucht. Dabei wurden jeweils signifikant längere Blickzeiten bei der Beobachtung von unmöglichen Ereignissen als Überraschung auf das beobachtete Ereignis interpretiert. So wurde bei 3- und 4-monatigen Babys aufgrund der Überraschungsreaktion auf das Durchdringen eines Objektes durch ein anderes Objekt auf ein erstes Verständnis der Objektpermanenz geschlossen (z. B. Baillargeon, Spelke & Wasserman, 1985; Baillargeon, 1987, 1993). Neuerdings wird diese Methode auch bei nicht-menschlichen Primaten angewendet (z. B. Hauser & Carey, 2003; Santos & Hauser, 2002; Santos, Miller & Hauser, 2003).

Bei der Überprüfung des Verständnisses eines komplexeren physikalischen Prinzips bei etwas älteren Kindern über die Unterscheidungsfähigkeit zwischen möglichen und unmöglichen Ereignissen stellt sich die Frage nach einer geeigneten abhängigen Variablen. Dabei hat sich Erfassung von Irritation oder Überraschung als problematisch erwiesen (vgl. S. 40). Relativ nahe liegend scheint es, die Unterscheidungsfähigkeit über die Vorhersage über das Eintreffen möglicher und unmöglicher Ereignisse zu erfassen.

Vorhersage möglicher und unmöglicher Transformationen

Mit der Erfassung der hypothetischen Voraussagen von möglichen und unmöglichen Transformationen wurde etwa das Verständnis biologischer Prozesse von Kindern untersucht (für einen Überblick siehe z. B. Goswami, 2001, S. 141ff.).

Rosengren, Kalish, Hickling und S. Gelman (1994, Studie 2; vgl. Rosengren, 1991) beispielsweise untersuchten mit dieser Methode das biologische Verständnis der Prozesse Wachstum und Metamorphose von 4- und 5-jährigen Kindern. Die Kinder hatten die hypothetischen Transformationen anhand von Bilderpaaren einer Ausgangslage sowie eines oder mehreren Endzuständen von Tieren zu beurteilen. Die Versuchsdurchgänge beinhalteten mögliche und unmögliche Transformationen: Mögliche Transformationen betreffend Wachstum bestanden in der Richtung von klein zu gross, diejenigen bei der Metamorphose in der Richtung von einfacher zu komplexer Form. Die umgekehrten Transformationsrichtungen waren biologisch unmöglich: betreffend Wachstum von gross zu klein und betreffend Metamorphose von komplex zu einfach. Die Kinder hatten die verschiedenen hypothetischen Verwandlungen danach einzuschätzen, ob sie möglich sind oder nicht. Damit konnte die Gruppe um Rosengren (1994) zeigen, dass bereits vierjährige Kinder die möglichen von den unmöglichen biologischen Veränderungsprozessen bei Tieren gut unterscheiden konnten. Im Alter von vier und fünf Jahren wussten die Kinder, dass ein Tier mit zunehmendem Alter nicht kleiner werden kann und sich Tiere nur von einer einfacheren zu einer komplexeren Form entwickeln können. Daraus wurde geschlossen, dass die Kinder im Kontext von biologischen Veränderungen ein frühes Verstehen von strengen Naturgesetzen zeigen. Die Erfassung der zusätzlichen Kommentare der Kinder ergab, dass die Kinder keine spontanen Bezüge zur Magie als möglicher Mechanismus machten.

Da die vier- und fünfjährigen Kinder die möglichen und unmöglichen biologischen Prozesse korrekt voraussagen konnten, verfügten sie in diesem Alter bereits über ein Verständnis biologischer Naturgesetze. Man kann annehmen, dass der Vorhersage von Möglichem und Unmöglichem bereits ein Identifizieren und Klassifizieren von Möglichem und Unmöglichem vorausgegangen ist. Erst dadurch kann das Kind eine Kategorie des Möglichen und damit ein physikalisches bzw. biologisches Weltbild aufbauen (Harris, 1994).

Schwierigkeiten verschiedener Art ergeben sich allerdings, wenn man das Verständnis eines Naturgesetztes – nicht mehr über das hypothetische Antizipieren, sondern – über die tatsächliche Präsentation unmöglicher (und möglicher) Ereignisse erfassen will.

Reaktionen auf präsentierte mögliche und unmögliche Transformationen

Will man das kindliche Realitätsverständnis über die Unterscheidungsfähigkeit zwischen präsentierten möglichen und unmöglichen Ereignissen erfassen, so muss die Präsentation der unmöglichen Ereignisse sehr natürlich und glaubwürdig erscheinen². Es liegt deshalb sehr nahe, auf Zaubertricks oder auf komplexe chemische oder physikalische Mechanismen zurückzugreifen, um die Durchbrechung der Naturgesetze überzeugend darzubieten. Dadurch verschiebt sich für die Versuchsperson aber auch die inhaltliche Bedeutung der Dichotomie möglich – unmöglich. Galt in den Aufgaben zur Vorhersage die Einschätzung des Normalfalls, so ist durch die überzeugende Präsentation eines unmöglichen Ereignisses – gerade dadurch, dass etwas an sich nicht Mögliches doch möglich erscheint – nicht mehr der Normalfall einzuschätzen. Wenn das Kind über das entsprechende Naturgesetz Bescheid weiss, wird es versuchen, das gesehene Unmögliche zu erklären. Die Einschätzung des Kindes, ob ein Ereignis möglich ist oder nicht, wird dabei nicht mehr nur das eingetroffene Ereignis auf der Basis rudimentärer Prinzipien umfassen, sondern auch die Verursachung des Ereignisses, die zur Durchbrechung des Prinzips geführt hat. In diesem Sinne wird ein erweitertes Realitätsverständnis, das Erklärungsversuche des unmöglichen Ereignisses beinhaltet, erfasst.

Dies zeigte sich deutlich nach Veränderung der Aufgabendarbietung der oben erwähnten Studie von Rosengren und Mitarbeitern (1994: Studien 2 und 3). Schätzten Kinder im Alter von 4 bis 6 Jahren die Wahrscheinlichkeiten des Eintreffens von möglichen und unmöglichen biologischen Veränderungen mit grosser Mehrheit richtig ein (Studie 2), änderte sich dies drastisch, wenn nach der Machbarkeit derselben biologischen Prozesse durch einen Zauberer gefragt wurde (Studie 3). Beim Zauberer wurden die unmöglichen Transformationen als genauso möglich eingestuft wie die normalen biologi-

² Beispielsweise präsentierte DeVries (1969) Verwandlungen von Tieren durch Aufsetzen von realistischen Tiermasken auf das Gesicht einer Katze. Gemäss ihren Ergebnissen hielten die untersuchten 3- und 4-jährigen Kinder die Transformationen einer Katze in einen Hund oder einen Hasen für möglich. Wie sich allerdings durch die Untersuchung von Hülksen, Sodian und Pickel (2001) gezeigt hatte, war die fehlende *theory of mind* bzw. Unterscheidungsfähigkeit zwischen Sein und Schein der Grund für dieses Resultat: Sie hatten die Verwandlung der Katze nicht als Verwandlung in ein anderes Tier verstanden, sondern sie gingen davon aus, dass die Katze komplett durch ein anderes Tier ersetzt wurde.

schen Prozesse. Im Gegensatz dazu schätzten die Erwachsenen auch die unmöglichen Prozesse als wenig wahrscheinlich ein, wenn sie ein Zauberer bewirken würde.

Die veränderten Einschätzungen sind einfach erklärbar: Bei der Studie 3 wurde nicht mehr nach dem alltäglichen Naturgesetz gefragt sondern nach der Verursachung durch eine speziell befähigte Person. Auch die Autoren interpretieren ihre Resultate dahingehend, dass die kindliche Kategorie des Möglichen beim Zauberer umfassender ist als in der Natur. Vieles, das für Erwachsene unmöglich ist, ist für Kinder (noch) möglich.

Allerdings können Kinder auch beim Zauberer durchaus verschiedene kausale Erklärungen in Erwägung ziehen. Kindliche Erklärungen zur Entstehung physikalisch unmöglicher Ereignisse umfassen mehrheitlich die beiden Kategorien „gezaubert mit Zaubertrick“ und „gezaubert mit richtiger Magie“, wie im folgenden Kapitel ausführlicher dargelegt wird.

2.3.2 Entwicklung des Verständnisses der Verursachung von gezauberten Ereignissen

Verschiedene Verursachungen gezauberter Ereignisse: Zaubertrick vs. richtige Magie

Kindliche Erklärungen zur Entstehung physikalisch unmöglicher Ereignisse umfassen mehrheitlich die beiden Kategorien „gezaubert mit Zaubertrick“ und „gezaubert mit richtiger Magie“. Erklärt das Kind das Ereignis mit richtiger Magie, so spricht man auch von magischer Kausalität oder magischem Denken. Unter magischem Denken wird in diesem Kontext der Glaube an die Verursachung eines Ereignisses durch Zauberkraft verstanden. Wie die folgenden Studien zeigten, sind für Kinder prinzipiell zwei Verursachungen unmöglicher Ereignisse denkbar: Entweder sie sind wirklich passiert oder sie sind getrickst.

In den Untersuchungen von Rosengren und Hickling (1994, Studien 1 und 2) wurden die Erklärungen vier- und fünfjähriger Kinder von so genannten gewöhnlichen, aussergewöhnlichen und gezauberten Transformationen erhoben. Unter „gewöhnlichen“ Transformationen verstanden die Autoren ganz alltägliche, auf einfachen physikalischen

Prinzipien beruhende Verwandlungen wie z. B. einer Schnur, die durch Zerschneiden in zwei kleinere Schnurstücke transformiert wurde. Bei den aussergewöhnlichen Transformationen wurden normale Objekte durch relativ komplexe und für Kinder kaum verständliche physikalische oder chemische Prozesse verändert. Beispielsweise veränderte sich die Farbe eines Spielzeugautos durch Eintauchen in heisses Wasser oder die farbige Flüssigkeit in einem Plastikröhrchen begann durch Knicken des Röhrchens zu leuchten. Als gezauberte Verwandlungen bezeichneten Rosengren und Hickling klassische Zaubertricks wie etwa die Verwandlung von zwei kurzen Schnurstücken in ein langes Schnurstück. Bevor den Kindern aber die Transformationen präsentiert wurden, mussten sie zuerst verschiedene Verwandlungsmöglichkeiten dieser Objekte vorschlagen und erklären, wie sie zu Stande kommen würden. Im Anschluss an die Transformationsvorschläge wurden die Objekte gewöhnlich, aussergewöhnlich oder durch zaubern verwandelt. Nach Beobachtung der Verwandlungen mussten die Kinder erklären, wie diese zu Stande gekommen waren.

Erwartungsgemäss beinhalteten die möglichen hypothetischen Verwandlungsvorschläge fast ausschliesslich physikalische Verwandlungen. Von sich aus kamen die Kinder selten auf die Idee, dass sich das Objekt auf gezauberte oder aussergewöhnliche Weise verändern könnte. Dieses Resultat deckt sich mit den Ergebnissen von anderen Studien (z. B. Huang, 1930; Rosengren et al., 1994; siehe weiter oben S. 41), bei denen kaum spontane Bezüge zur Magie gemacht wurden. Die Erklärungen nach Beobachtung der verschiedenen Transformationen wurden vorwiegend eingeteilt in die drei Kategorien „physikalisch/natürlich“ (normaler physikalischer Vorgang), „getrickst“ (die Verwandlung wurde nur vorgetäuscht) oder „magisch“ (die Verwandlung war echt und durch echte Magie entstanden). Es zeigte sich, dass die gewöhnlichen Verwandlungen ausschliesslich mit einem physikalischen Mechanismus erklärt wurden. Die gezauberten Transformationen hingegen waren bei den Vierjährigen vorwiegend durch richtige Magie und bei den Fünfjährigen durch Zaubertricks entstanden. Die Erklärungen zu den aussergewöhnlichen Verwandlungen waren bei den 4-jährigen ähnlich wie bei den gezauberten Verwandlungen, bei den 5-jährigen waren sie mehrheitlich physikalischer Natur.

Damit wurde gezeigt, dass schon 4- bis 5-jährige Kinder verschiedene Transformationen den Kategorien „physikalisch möglich“ und „gezaubert“ (die beiden Kategorien

„Tricks“ und „Magie“ zusammen genommen) gut und klar zuordnen können. Allerdings hielten die Vierjährigen – im Gegensatz zu den Fünfjährigen – die Magie für einen plausiblen Mechanismus von an sich unmöglichen und aussergewöhnlichen Transformationen. Man kann sagen, dass Kinder im Alter von 4 bis 5 Jahren ein umfassendes Wissen über die Kategorie „Zaubern“ haben, sie sich aber in ihren Überzeugungen über die Natur des „Zauberns“ oder der Magie unterscheiden.

Der Erklärungsversuch von Rosengren und Hickling (1994, p. 75), dass die Kinder, wenn sie die kausalen Erklärungen nicht kennen, auf die Kategorie „Magie“ zurückgreifen, scheint plausibel. Sie vermuten, dass die vierjährigen Kinder die an sich unmöglichen Transformationen, die ausserhalb ihres Erfahrungsbereiches lagen, vorschnell als tatsächliche Ereignisse akzeptierten. Den Fünfjährigen fehlte vermutlich auch die Einsicht in die Tricks, dennoch wagten sie den Realitätscharakter der Transformationen anzuzweifeln.

In sehr ähnlicher Weise interpretierten Chandler & Lalonde (1994) die magischen Erklärungen drei- und vierjähriger Kinder bei ihrem ersten Experiment. Sie benutzten die gleiche Apparatur wie sie Baillargeon zur Untersuchung der Objektpermanenz benutzt hatte (z. B. Baillargeon et al., 1985; Baillargeon, 1993), bei welcher ein festes Objekt (Schirm) ein anderes Objekt (Holzklotz) scheinbar durchdringt. Anstelle von 3- und 4-monatigen Kindern wie bei Baillargeon wurden in dieser Studie 3- und 4-jährige Kinder untersucht. Auch wurde nicht die Überraschungsreaktion mit Blickzeitenmessung auf die Beobachtung des unmöglichen Ereignisses erfasst, sondern die Kinder wurden explorativ untersucht. Die spontanen Reaktionen und Beschreibungen dessen, was sie beobachtet hatten, wurden notiert und in die drei Kategorien „Magie“, „Trick“ und „Realität“ (= Kategorie des Möglichen) eingeteilt. Es zeigte sich, dass die Kinder zunächst mehrheitlich das Durchdringen eines Objektes durch ein anderes Objekt als ein Resultat von echter Magie erklärten. Das unmögliche Ereignis wurde in diesem Experiment aber zweimal wiederholt, um allfällige Veränderungen in den kindlichen Erklärungen festzustellen. Es zeigte sich, dass die Kinder nach der zweiten Beobachtung immer noch häufig „Magie“-Antworten gaben, zunehmend aber auch vermuteten, es handle sich um einen Trick. Nach der dritten Präsentation durften Kinder die Apparatur selber untersuchen. Nach Inspektion der Trickapparatur war fast allen Kindern klar geworden, dass es sich dabei um einen Trick gehandelt hatte.

Auch in dieser Studie wurde gezeigt, dass jüngere Kinder unmögliche Ereignisse als gezauberte Ereignisse einschätzen, sie sich aber im Verständnis von Zauberei unterscheiden. Gleichzeitig wurde deutlich, wie instabil die magischen Erklärungen nach wiederholter Beobachtung beim gleichen Ereignis und v. a. nach Inspektion der Apparatur waren.

Chandler und Lalonde (1994) legten die magische Kategorie im Hinblick auf ihre Funktion als eine Art Restkategorie oder eine spezifische „Wirklichkeitskategorie“ aus. Diese komme immer dann zum Zuge, wenn den Kindern Erklärungen zu an sich unmöglichen Ereignissen fehlen. So könnte die Kategorie der Magie der Erhaltung des physikalischen Weltbildes dienen, indem das Unmögliche seine Berechtigung erhält, ohne die physikalischen Prinzipien infrage stellen zu müssen. Auch das zweite Experiment von Chandler und Lalonde (1994; vgl. S. 36ff.), bei dem sich gezeigt hatte, dass ein durch ein Zauberer verursachtes Ereignis das physikalische Verständnis viel weniger gefährdet, als wenn es durch „glaubwürdigere“ Akteure wie ein Priester oder Psychologe verursacht wird, kann auf die gleiche Weise interpretiert werden. Ob diese Ereignisse von den Kindern als tatsächliche Realität (richtige Magie) oder als Täuschung (Tricks) verstanden werden, ist damit aber noch nicht geklärt. Chandler und Lalonde (1994) gehen nicht davon aus, dass die Kategorie des Magischen bei jungen Kindern verschwindet, sondern sie vermuten eher, dass sie schnell lernen würden, dass diese Kategorie in unserer rationalen Welt keinen Platz habe, weshalb die Kinder diese Kategorie auch nicht mehr erwähnen würden. Gleichzeitig wird ein Teil der noch unverständlichen Ereignisse aber durch zunehmendes Wissen aufgeheilt, weshalb zur Erklärung von unmöglichen Ereignissen keine Magie mehr herhalten muss. Dieser individuelle Entwicklungsprozess wurde auch in der Entwicklung des Weltwissens aus phylogenetischer Perspektive beobachtet (vgl. Schmid, 1998).

Spontane Bezüge zur Magie, die ein hypothetisches unmögliches Ereignis verursacht haben könnte, werden kaum gemacht. Dennoch ist diese Rationalität limitiert und zwar dort, wo bestimmten Akteuren besondere Fähigkeiten zugeschrieben werden. Vielleicht könnte man sagen, dass die Limitierung erst in einem erweiterten Kausalitätsbereich beobachtet werden kann: Werden nicht nur die Erklärungsversuche der Ereignisse erfasst, sondern werden auch verschiedene Verursachungen in Erwägung gezogen, unterscheiden sich die Kinder deutlich.

Unterschiede in der Verursachung gezauberter Ereignisse durch verschiedene Akteure

Die Einschätzung von Möglichkeit bzw. Unmöglichkeit eines Ereignisses hängt also wesentlich davon ab, wie es verursacht wird. Es sind Umstände denkbar, in denen auch ein an sich unmögliches Ereignis doch möglich wird. Solche Umstände sind beispielsweise gegeben, wenn das Ereignis von einem Zauberer verursacht wird.

In der bereits auf S. 43f. beschriebenen Studie von Rosengren & Hickling (1994) mussten die 4- und 5-jährigen Kinder die vorgeführten möglichen und unmöglichen Transformationen nach ihrer Machbarkeit durch verschiedene Akteure einschätzen. Präsentiert wurden neben ganz gewöhnlichen physikalischen Vorgängen (wie Zerschneiden einer Schnur) auch aussergewöhnliche auf komplexer chemischer oder physikalischer Basis beruhende Transformationen (wie etwa ein Stab mit farbiger Flüssigkeit, die nach Knicken des Stabes zu leuchten beginnt) sowie klassische Zaubertricks (sog. gezauberte Transformationen). Diese Transformationen mussten von den Kindern nach ihrer Machbarkeit durch sie selber, einen Freund, Vater oder Mutter, ihre Lehrperson und einen Zauberer eingeschätzt werden. Die Ergebnisse zeigten, dass die Kinder die verschiedenen Akteure insgesamt für deutlich befähigter hielten, die gewöhnlichen als die aussergewöhnlichen und gezauberten Transformationen zu Stande zu bringen. Die Kinder unterschieden demnach die auf einfachen physikalischen Prinzipien beruhenden Transformationen von solchen, die auf Mechanismen beruhten, die den Kindern nicht unmittelbar verständlich waren auch bei der Machbarkeitseinschätzung. Während bei den gewöhnlichen Verwandlungen die Akteure als gleich hoch befähigt beurteilt wurden, zeigten sich bei den aussergewöhnlichen und gezauberten Verwandlungen Unterschiede in den Zuschreibungen zu verschiedenen Akteuren. Bei den gezauberten Verwandlungen galten die Zauberer als die fähigsten, während die Kinder sich selbst als am wenigsten befähigt einschätzten. Die Lehrperson wird nur als leicht fähiger, als sie sich selber sahen, beurteilt. Bei den aussergewöhnlichen Transformationen hingegen wurde die Fähigkeit nicht mehr nur dem Zauberer zugeschrieben sondern auch den anderen Akteuren, dies aber nicht so ausgeprägt wie beim Zauberer. Aussergewöhnliche Transformationen galten im Durchschnitt als machbarer als die gezauberten Ereignisse. Insbesondere bei der Einschätzung der eigenen Machbarkeit waren die Kinder bei den gezauberten Transformationen tiefer als bei den aussergewöhnlichen. Zusammengefasst zeigte sich also, dass die verschiedenen Transformationen hinsichtlich ihrer Produzier-

barkeit durch verschiedene Akteure in der folgenden Reihenfolge unterschieden wurden: gewöhnliche > aussergewöhnliche > gezauberte. Zudem galt der Zauberer als ein Akteur, der über die höchste Fähigkeit verfügt, da er im Unterschied zu den anderen Transformationen sowohl bei den gezauberten als auch bei den aussergewöhnlichen höher eingeschätzt wurde als die anderen Akteure.

Bemerkenswert fanden die Autoren u. a. die unterschiedlichen Beurteilungen der eigenen Fähigkeit zwischen den aussergewöhnlichen und den gezauberten Ereignissen, die gesehenen Ereignisse selber zu produzieren. Wieso trauten sich die Kinder die aussergewöhnlichen Ereignisse eher zu als die gezauberten? Da sich die Transformationen beider Arten ziemlich abrupt ereigneten, konnte es nicht an der zeitlichen Handlungsabfolge der Transformation liegen. Zudem achteten Rosengren und Hickling besonders darauf, dass bei beiden Arten von Transformationen ähnlich diskrete Handlungen (z. B. ziehende, stossende Handlungen) zur Verursachung der Transformationen vollzogen wurden, so dass sie ausschliessen konnten, dass die Kinder subtile Differenzen in der Art der Handlungen bemerkt haben könnten. Sie hielten es auch für unwahrscheinlich, dass die Kinder mehr Kenntnisse über die „Tricks“ der aussergewöhnlichen Transformationen als über diejenigen der gezauberten Ereignisse hatten.

Betrachtet man die unterschiedlichen Machbarkeitszuschreibungen insbesondere der Akteure Zauberer und Kind selbst, so kann man diese auch im Lichte des erweiterten Realitätsverständnisses betrachten. Geht man davon aus, dass bei der eigenen Zuschreibung eingeschätzt wird, was im Normalfall möglich ist und beim Zauberer, nur was in Ausnahmefällen möglich ist, so wären diese Einschätzungen der Kinder schon ziemlich realistisch. Aber wie wissen die Kinder, dass die aussergewöhnlichen auch ohne spezifische Fingerfertigkeit oder magische Fähigkeiten zu Stande kommen? Ein Hinweis kommt aus der Studie selbst, obwohl die Autoren hier keinen Zusammenhang machten. Vor der Studie wurde in einem Vortest mit Erwachsenen die Vertrautheit der verschiedenen Transformationen erfasst. Es zeigte sich folgende Reihenfolge mit abnehmender Vertrautheit: gewöhnliche > aussergewöhnliche > gezauberte Transformationen. Es wäre also durchaus möglich, dass sich die Kinder die aussergewöhnlichen Transformationen mehr zutrauten, weil sie ihnen aus dem Alltag bereits vertrauter waren als die Zaubertricks. Ereignisse, die sie aus dem Alltag kennen, funktionieren nach physikalischen Gesetzen, weshalb sich die Kinder die aussergewöhnlichen Ereignisse aufgrund der hö-

heren Vertrautheit mehr zugetraut haben könnten. Hierfür sprechen auch die kindlichen Vermutungen der Fünfjährigen zur Entstehung der Transformationen: Diese erklärten die gezauberten Transformationen mehrheitlich mit Trick, während die aussergewöhnlichen mehrheitlich physikalisch/natürlich erklärt wurden. Man kann sich auch fragen, ob mit den Machbarkeitszuschreibungen ein implizites Kausalitätsverständnis erfasst wurde, da der einzige Unterschied zwischen den beiden Ereignissen auf den unterschiedlichen Entstehungsmechanismen beruht: Aussergewöhnliche Transformationen funktionieren nach Naturgesetzen (auch wenn diese kompliziert sind) im Gegensatz zu den Zaubertricks. Es scheint, als würde mit der eigenen Zuschreibung das Verständnis des Naturgesetzes (vermutlich das implizite) erhoben und mit der Zuschreibung zum Zauberer eine andere Realitätsdimension. Damit ist die Erfassung mit der Machbarkeitszuschreibung zu verschiedenen Akteuren der direkten Erfragung überlegen, weil den Kindern das Wissen über die Verursachung noch lange fehlt und ihre Verbalisierungsfähigkeit lange eingeschränkt ist.

3 Empirische Untersuchung

Die empirische Untersuchung gliederte sich in die drei Teile „Prädiktion der Farbmischungsresultate“, „*magic experiment*“ sowie „Kausalitätsverständnis gezauberter Ereignisse“. Die Untersuchung aller drei Teile fand bei den gleichen Versuchspersonen und in derselben Sitzung statt. Bevor die Ergebnisse der drei Phasen in einzelnen Kapiteln (Kapitel 3.2 bis 3.4) ausführlich dargestellt und diskutiert werden, werden im ersten Kapitel zur Methode die gemeinsamen Grundlagen aller drei Experimentalphasen dargestellt: Beschreibung der Stichprobe, Versuchsmaterial und –durchführung sowie die Versuchspläne der einzelnen Experimentalphasen.

3.1 Methode

Die vorliegende Untersuchung bestand aus einem Experiment mit drei Phasen: Prädiktion der Farbmischungsresultate, *magic experiment* und Kausalitätsverständnis gezauberter Ereignisse. Einführend werden die drei Phasen kurz umrissen.

Phase 1: Prädiktion der Farbmischungsresultate

Den Versuchspersonen wurden verschiedene Gläserpaare mit gleich farbigen aber verschieden hellen Flüssigkeiten dargeboten. Die Kombination der Farbintensität der beiden Ursprungsflüssigkeiten wurde variiert: zwei gleich helle Flüssigkeiten (gleiche Intensitätskombination) und zwei unterschiedlich helle Flüssigkeiten (unterschiedliche Intensitätskombination). Die Versuchspersonen hatten das Mischungsresultat der beiden Ursprungsflüssigkeiten hinsichtlich der Helligkeit vorausszusagen. Das prädiktive Mischungsurteil wurde pro Intensitätskombination zweimal erhoben, um die Urteile konsistenten Mischungsurteilen (insbesondere Durchschnitts-, Additionsregel und inkonsistente Regel) zuweisen zu können. Im Anschluss an das jeweilige Mischungsurteil sollte die Versuchsperson die subjektive Sicherheit ihres Urteils anhand einer Skala beurteilen.

Phase 2: *Magic experiment*

In einer zweiten Phase wurden die Flüssigkeiten der zwei Ursprungsgläser vor den Augen der Versuchspersonen zusammengeleert. Allerdings waren bei jeder Intensitätskombination (gleiche und unterschiedliche Intensitätskombinationen) die Ursprungsflüssigkeiten je einmal chemisch manipuliert worden, so dass sich pro Intensitätskombination je zwei verschiedene Mischungsresultate ergaben:

- a) physikalisch mögliche bzw. korrekte Mischungsresultate, welche der korrekten Durchschnittsregel entsprachen,
- b) physikalisch unmögliche bzw. unkorrekte Mischungsresultate, welche dunkler waren als die beiden Ursprungsflüssigkeiten und der Additionsregel entsprachen.

Dadurch wurden sowohl den korrekten Durchschnittsregelanwendern als auch den Additionsregelanwendern je ein erwartetes und ein unerwartetes Mischungsresultat pro Intensitätskombination vorgeführt.

Die Reaktionen auf die präsentierten Mischungen wurden einerseits mit einer Zauberskala erfasst, anhand derer der Realitätscharakter der Mischungen eingeschätzt werden musste. Andererseits wurde aber auch nach dem bewussten Umgang mit der prädiktiven Regel nach Beobachtung der Mischungsvorgänge mit einer subjektiven Sicherheitsskala gefragt.

Phase 3: Kausalitätsverständnis gezauberter Ereignisse

Die Stimuli der Phase 3 waren dieselben wie in Phase 2. Anhand der vier Mischungsresultate sollte die Versuchsperson abschätzen, ob nebst der Versuchsleiterin auch andere Personen zu den gleichen Mischungsresultaten befähigt sind. Die Versuchspersonen sollten bei jedem Versuchsdurchgang Machbarkeitszuschreibungen zu den vier Akteuren anderer Zauberer, Lehrperson, Eltern der Versuchsperson und Versuchsperson selbst machen.

3.1.1 Versuchspersonen

An der vorliegenden Untersuchung nahmen insgesamt 172 Versuchspersonen aus sechs Altersgruppen teil. Die Kinder stammten aus zwei verschiedenen Schulhäusern in Zürich. Das eine Schulhaus befand sich in einem Quartier mit einem für die Stadt Zürich mittleren Ausländeranteil und das andere gehörte zu einem Quartier mit überwiegend mittelständischen Einwohnern. Die erwachsenen Versuchspersonen waren Psychologiestudierende an der Universität Zürich, welche sich zum Erhebungszeitpunkt mehrheitlich im dritten Semester des Studiums befanden. Von insgesamt sechs Kindern konnten die Daten nicht verwendet werden, da bei vier Kindergartenkindern und zwei Zweitklässlern das Experiment nicht korrekt durchgeführt werden konnte, weil sie mangels Deutschkenntnissen die Instruktion nicht verstanden hatten oder die vier Gläser der Farbskala nicht richtig ordnen konnten.

Allerdings wurde die Stichprobe für die dargestellten Ergebnisse nochmals reduziert³. Die Gruppe der Drittklässler bzw. Neunjährigen (18 Kinder) wurde weggelassen, um die Altersabstände zwischen den übrigen Altersgruppen konstant zu halten. Es war ursprünglich auch nicht beabsichtigt gewesen, Daten von Neunjährigen zu erfassen, da aber in der Kommunikation mit der Lehrerin ein Fehler unterlief, welcher erst während der Versuchsdurchführung bemerkt wurde, wurde das Experiment mit allen Kindern, welche daran teilnehmen wollten, trotzdem durchgeführt. Zudem wurden die Altersgruppen auf eine einheitliche Stichprobengrösse von 30 Kindern bzw. 20 Erwachsenen reduziert.

Die gesamte Stichprobe, die den Auswertungen zu Grunde lag, bestand aus 30 Kindergartenkindern (Durchschnittsalter: 6;1⁴), 30 Zweitklässlern (Durchschnittsalter: 8;1), 30 Viertklässlern (Durchschnittsalter: 10;1), 30 Sechstklässlern (Durchschnittsalter: 12;7) und 20 Erwachsenen (Durchschnittsalter: 25;10). In der Folge werden in dieser Arbeit die Gruppen der Kinder nicht mehr mit Bezug auf die betreffende Schulklasse sondern mit Bezug auf das durchschnittliche Alter der jeweiligen Gruppe bezeichnet. So wird

³ Die ursprüngliche Stichprobe bestand aus 30 Kindergartenkindern, 31 Zweit-, 18 Dritt-, 35 Viert- und 30 Sechstklässlern sowie 21 Erwachsenen.

⁴ Die Altersangabe erfolgt auf Monate genau. Vor dem Semikolon befindet sich das Alter in Jahren, hinter dem Semikolon befinden sich die zusätzlichen Monate.

nur noch von den Sechs-, Acht-, Zehn- und Zwölfjährigen sowie den Erwachsenen die Rede sein. Die Geschlechterverteilung in den Altersgruppen war mit Ausnahme der Sechstklässler ziemlich ausbalanciert. In der Altersgruppe der Sechstklässler befanden sich zwanzig Mädchen und zehn Knaben. Die vollständige Übersicht über die Alters- und Geschlechtsangaben der fünf Altersgruppen ist der Tabelle 3.1.1 zu entnehmen.

Tabelle 3.1.1

Übersicht über die Verteilungen der auf der Datenauswertung basierenden Versuchspersonen pro Altersgruppe

Altersgruppe	n	Durchschnittliches Alter	Altersbereich	Geschlechterverteilung (f / m)
Sechsjährige	30	6;01	05;01 – 07;01	17 / 13
Achtjährige	30	8;01	06;11 – 09;01	14 / 16
Zehnjährige	30	10;01	09;02 – 11;05	18 / 12
Zwölfjährige	30	12;07	11;04 – 13;07	20 / 10
Erwachsene	20	25;10	19;07 – 36;09	12 / 08

Anmerkung. Die Altersangaben beziehen sich auf das Jahr (vor dem Semikolon) und die Monate (nach dem Semikolon).

3.1.2 Versuchsmaterial

Das Versuchsmaterial bestand in erster Linie aus den zu mischenden, paarigen Ursprungsflüssigkeiten sowie einer vierstufigen Farbintensitätsskala, bestehend aus vier verschiedenen Flüssigkeiten. Alle farbigen Flüssigkeiten wurden in quasi-zylindrischen Joghurtgläsern mit einem Fassungsvermögen von ca. 180 ml, welche durch Plastikdeckel verschliessbar waren, dargeboten. Die vier Gläser der Farbskala enthielten 100 ml Flüssigkeit, die Gläser mit den Ursprungsflüssigkeiten enthielten je die Hälfte, d. h. 50 ml. Die farbigen Flüssigkeiten bestanden aus Wasser und zugegebener hochkonzentrierter flüssiger rotvioletter Wasserfarbe (Marke „Ecoline“, Nr. 545).

Farbintensitätsskala und Ursprungsflüssigkeiten für physikalisch korrekte Mischungsergebnisse

Wie der Abbildung 3.1.1 zu entnehmen ist, bestand die Farbintensitätsskala aus vier nebeneinanderstehenden Gläsern, deren Flüssigkeiten sich hinsichtlich der Farbintensi-

tät von Glas zu Glas kontinuierlich unterschieden (vgl. auch Abbildung der Skala des Experiments im Anhang B, S. 275). Dabei stellte Stufe bzw. Helligkeit 1 die hellste Flüssigkeit und Stufe bzw. Helligkeit 4 die dunkelste Flüssigkeit dar.

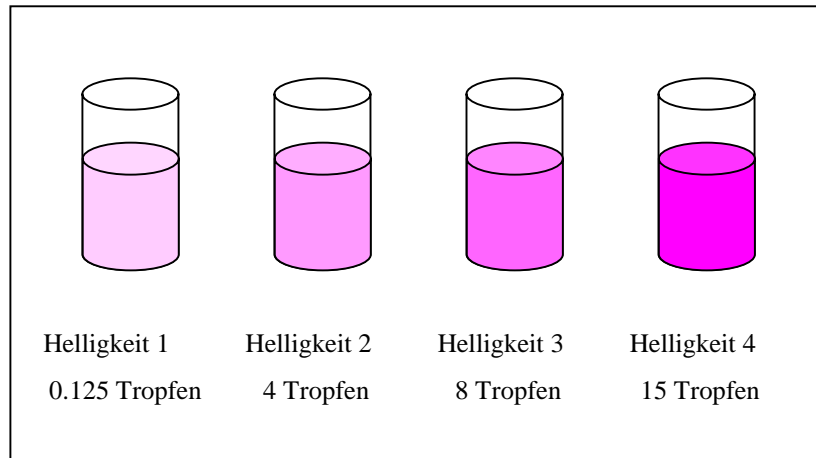


Abbildung 3.1.1. Schematische Darstellung der Farbtintensitätsskala mit vier Helligkeitsstufen mit Angabe der Farbstoffmenge (Anzahl Tropfen aus einer 1ml-Einwegspritze) pro 100 ml Wasser.

Eine wichtige Bedingung für die Herstellung der Farbskala war, dass die Intensitätsunterschiede der Rotviolettöne der vierstufigen Farbskala gut erkennbar sein sollten. Damit der wahrgenommene Farbtintensitätsanstieg zwischen den Gläsern in gleich grossen Sprüngen erfolgte, musste die zugegebene Menge an Farbkonzentration exponentiell ansteigen. Dieser funktionale Zusammenhang wird durch die Stevens'sche Potenzfunktion beschrieben. Da die vierstufige Skala aber nicht nur der Durchführung des Diskriminationstests diente, sondern gleichzeitig auch den Ursprungsflüssigkeiten und den Mischungsresultaten entsprechen sollte, musste auch berücksichtigt werden, dass sich aus der Mischung von Helligkeit 1 und 3 die Helligkeit 2 ergab. Bei der Bestimmung der Helligkeiten 1 bis 3 musste also darauf Rücksicht genommen werden, dass das Mischungsresultat der Kombination von Helligkeit 1 und 3 nach der tatsächlichen Mischung der Farbtintensität der Helligkeit 2 entsprach. Dies wurde mit den in Abbildung 3.1.1 angegebenen Wasserfarbvolumina erreicht. Konkret wurden bei der Mischung von Ursprungsflüssigkeit 1 und 3, mit je einem Volumen von 50 ml, 0.0625 (Flüssigkeit 1) und 4 Tropfen Wasserfarbe (Flüssigkeit 3) zusammengeleert, was zu einem Mischungsresultat von 100 ml Wasser und 4.0625 Tropfen Wasserfarbe führte. Dieses Mischungsresultat entsprach wahrnehmungsmässig der Helligkeit 2 der oben dargestellten Farbskala. Helligkeit 4 schliesslich wurde so bestimmt, dass sie von mehreren Vorversuchspersonen hinreichend gut von Stufe 3 zu unterscheiden war. Da die Helligkeit 4 dem

erwarteten Resultat von Anwendern der Additionsregel möglichst gut entsprechen sollte, durfte es nicht zu dunkel gewählt werden.

Ursprungsflüssigkeiten für physikalisch unkorrekte Mischungsresultate

Die physikalisch unkorrekten Mischungsresultate wurden mithilfe eines Basenindikators erzielt. Hierzu wurde der Basenindikator Phenolphthalein gewählt, der in saurem und neutralem Milieu farblos ist, in basischem Milieu – der Umschlagsbereich liegt bei den pH-Werten 8.2 bis 10.0 – aber rotviolett sichtbar wird (und deshalb ja auch als Indikator für basisches Milieu eingesetzt wird). Giesst man also in eine neutrale – z. B. wässrige – Flüssigkeit, die Phenolphthalein enthält, eine basische Flüssigkeit wie etwa Natronlauge, so verschiebt sich der pH-Wert der ursprünglich neutralen Flüssigkeit in den basischen Bereich und der Indikator wird sichtbar. Der Indikator verhält sich im farbigen sichtbaren Bereich wie jeder andere Farbstoff: Seine Helligkeit bzw. Farbintensität ist von der Menge bzw. der Konzentration des Indikators abhängig. Somit kann die Farbintensität der Indikatorlösung variiert und exakt, entsprechend der vierstufigen Skala, gewählt werden. Theoretisch könnten also einzig durch Variation der Indikatorkonzentrationen alle vier Farbintensitäten der Skala von Abbildung 3.1.1 alleine mit Phenolphthalein in basischem Milieu erreicht werden. Auch aus der Mischung von zwei farblosen manipulierten Flüssigkeiten – eine enthält den Indikator und ist neutral, die andere ist basisch – können durch Variation der Indikatormenge verschieden intensive farbige Flüssigkeiten erzeugt werden. Genau dieses Prinzip wurde für den Probeversuchsdurchgang genutzt. Für die Versuchsdurchgänge mussten nur noch die manipulierten Ursprungsflüssigkeiten mit einem Farbstoff, der den gleichen Farbton wie der Indikator hat, angefärbt werden. Die Mengen des Indikators wurden so gewählt, dass sich aus der Mischung der Helligkeit 2 und 2 (gleiche Intensitätskombination) sowie Helligkeit 1 und 3 (unterschiedliche Intensitätskombination) je ein Mischungsresultat der Helligkeit 4 ergab. Mit der Wasserfarbe „Ecoline“ (Nr. 545) wurde ein Farbstoff gefunden, der wahrnehmungsmässig der Indikatorfarbe von Phenolphthalein entsprach. Ein weiterer Vorteil dieses Indikators bestand darin, dass die Verfärbung durch Mischung mit basischer Flüssigkeit etwa mit derselben Geschwindigkeit erfolgte, wie sich der normale Farbstoff durch Vermischung gefärbter Flüssigkeiten auf das gesamte Volumen verteilt. Die Farbintensivierung präsentierte sich dadurch auf sehr natürliche Art und Weise. Es war damit unmöglich, anhand des Mischungsvorganges alleine zu merken, wann manipuliert wurde und wann nicht.

Bei der Bereitstellung der Ursprungsgläser mussten also zwei der vier Paare von Ursprungsflüssigkeiten so präpariert werden, dass je eine angefärbte Flüssigkeit mit unsichtbarem Indikator und je eine mit der unsichtbaren Base versehen war, damit der Indikator durch Mischung mit einer basischen Flüssigkeit auch entsprechend reagierte und die Mischung dunkler wurde.

Anleitung zur Erstellung der verwendeten Ursprungsflüssigkeiten

Da sich Phenolphthalein nur in Alkohol löst, wurde eine Indikator-Lösung erstellt, bei der 0.5g Indikator in eine 1l-Flasche Brennsprit zugegeben wurde. Diese Lösung wurde noch stark mit Wasser verdünnt (siehe Tabelle 3.1.2), so dass der Geruch des Alkohols für die Versuchspersonen nicht wahrnehmbar war. Als basische Flüssigkeit wurde eine Natronaugenlösung hergestellt, die pro Liter H_2O 20 Plättchen fester Natronlauge (NaOH) enthielt. Durch Schütteln der Flüssigkeit lösten sich die Plättchen schnell im Wasser auf. Allerdings wurde die Natronaugenlösung nach einigen Tagen etwas trüb, weshalb den anderen Lösungen (siehe Tabelle unten) Spuren von Mehl beigegeben werden musste. Die im Fachhandel erhältliche flüssige Wasserfarbe „Ecoline“ (Nr. 545) war so hoch konzentriert, dass nur wenige Tropfen zur Anfärbung von Wasser ausreichten. Um die Tropfen gut dosieren zu können, wurde eine 1ml-Einwegspritze verwendet. 18 Tropfen Wasserfarbe aus dieser Spritze entsprachen 1 ml der gekauften flüssigen Wasserfarbe.

Tabelle 3.1.2

Übersicht über die Zusammensetzungen der Ursprungsflüssigkeiten mit einem Volumen von 50 ml bzw. 200ml

Lösung	Indikatorlösung	H_2O	NaOH-Lösung	Anzahl Tropfen Wasserfarbe
H_2O -Indikator	Nur wenig zugeben	Auffüllen auf 50 ml Volumen		
H_2O -NaOH			50 ml	
1-NaOH			200 ml	1
2-NaOH			50 ml	2
2-Indikator	5 ml	45ml		2

3-Indikator	5 ml	45ml	4
2-normal		50ml	2
3-normal		50ml	4

Anmerkung. Bei der Lösung „1-NaOH“ bezieht sich die Anzahl Tropfen auf ein Volumen von 200 ml. Die Tropfen der Wasserfarbe wurden aus einer Spritze von einem Volumen von 1 ml abgezählt. Die Zahlen bei den Bezeichnungen der verschiedenen Lösungen weisen auf die Helligkeitsstufen der verwendeten Farbintensitätsskala.

Die dargestellten Lösungen wurden für folgende Versuchsdurchgänge verwendet:

- *Probeversuchsdurchgang:* „H₂O-Indikator“ und „H₂O-NaOH“,
- *Gleiche Intensitätskombination mit physikalisch korrektem Mischungsresultat:* „2-normal“ und „2-normal“,
- *Gleiche Intensitätskombination mit physikalisch unkorrektem Mischungsresultat:* „2-Indikator“ und „2-NaOH“,
- *Unterschiedliche Intensitätskombination mit physikalisch korrektem Mischungsresultat:* „1-NaOH“ (da die Natronlaugenlösung sich genauso gut mit Wasser vermischte wie eine normale Wasserlösung wurde auf eine zusätzliche Mischung von „1-normal“ verzichtet) und „3-normal“,
- *Unterschiedliche Intensitätskombination mit physikalisch unkorrektem Mischungsresultat:* „1-NaOH“ und „3-Indikator“.

Ausserdem wurde ein in weisses Papier gekleidetes ca. 15 cm hohes Podest mit einer 20 cm hohen weissen Hinterwand zur Hilfe genommen, welches auf den Tisch vor die Versuchsperson gestellt wurde. Bei der Versuchsdurchführung wurden die Gläser pro Versuchsdurchgang auf das Podest gestellt, wodurch sämtliche Stimuli vor derselben weissen Hinterwand des Podests präsentiert wurden. Damit sollte zudem gewährleistet sein, dass alle Versuchspersonen bei der Beurteilung der Farbintensitäten die gleichen Bedingungen hatten. Gleichzeitig wurden die Stimuli damit vor einem für die Intensitätsdiskriminierung idealen Hintergrund präsentiert. Ein weiterer Vorteil dieses Podests bestand darin, dass sich die Gläser durch die höhere Position auf der durchschnittlichen Augenhöhe der Kinder befanden und somit ebenfalls gut wahrnehmbar waren.

3.1.3 Versuchsdurchführung

Das Experiment wurde in Einzelsitzungen mit den Versuchspersonen durchgeführt. Die Durchführung des Experiments dauerte pro Versuchsperson durchschnittlich 20 Minuten. Bei der Versuchsdurchführung war neben der Versuchsperson und der Versuchsleiterin auch eine studentische Hilfskraft anwesend, welche die Gläser vorbereitete und damit einen kontinuierlichen Ablauf der Experimente gewährleistete.

Die Untersuchungen fanden für die Altersgruppe der Kindergartenkinder in ihrem Kindergarten statt. Alle anderen Kinder wurden in der jeweiligen Schule in einem separaten Raum untersucht. Die Untersuchung der Erwachsenen fand in den Räumen des Psychologischen Institutes der Universität statt. Bei der Wahl des Durchführungsortes wurde darauf geachtet, dass der Raum möglichst viel Tageslicht zuliess, damit die Farbintensitätsunterschiede gut wahrnehmbar waren.

Die Versuchsperson sass an einem Tisch, auf dem sich alle vorbereiteten Gläser für die Versuchsdurchführung befanden. Zuerst mussten die Versuchspersonen einen Diskriminations- und Seriationstest durchlaufen, indem sie die vier Gläser der Skala in die richtige Reihenfolge nach aufsteigender oder absteigender Helligkeit bringen mussten.

Anschliessend wurden zunächst alle Gläserpaare des Probeversuchsdurchganges und der vier Experimentaldurchgänge gezeigt und das prädiktive Mischungsurteil und die subjektive Sicherheit dieses Urteils (Ratingskala 1, siehe Anhang D, S. 281) erfasst. Die Erhebung der angewendeten Mischungsregel musste vor dem Mischen durchgeführt werden, da die Rückmeldung durch die Mischungsergebnisse die Antworten beeinflusst hätte. In einem zweiten Teil wurden die Flüssigkeiten - in derselben Reihenfolge wie vorher - durch die Versuchsleiterin dann tatsächlich gemischt. Nach der Präsentation der Mischungsvorgänge wurden pro Versuchsdurchgang die weiteren abhängigen Variablen in der folgenden Reihenfolge erfasst: Einschätzung des Mischungsergebnisses, Realitätseinschätzung mit der Zauberskala (Ratingskala 2, siehe Anhang D, S. 282), subjektive Sicherheit der prädiktiven Mischungsregel nach der Mischung (Ratingskala 3, siehe Anhang D, S. 283) und die Einschätzungen der Machbarkeitszuschreibungen zu verschiedenen Akteuren.

Um die Versuchsperson auf das Thema „Zaubern und Farbmischungstrick“ einzustellen, wurden ihr eingangs die Gläser mit farblosen Flüssigkeiten präsentiert. Die Versuchsperson wurde gefragt, was mit der einen Flüssigkeit geschehen werde, wenn man die Flüssigkeit der anderen Komponente dazu schüttet: Würde es heller oder dunkler werden? Würde es gleich bleiben? Danach wurden diese Gläser zur Seite gestellt und die Gläser mit des nächsten Versuchsdurchganges aufgebaut. Erst nachdem alle Versuchsdurchgänge präsentiert wurden, wurden die Mischungen auch tatsächlich zusammen geleert. Die genaue Instruktion ist dem Anhang A (S. 268ff.) zu entnehmen.

3.1.4 Versuchsplan

Da das Untersuchungsdesign in den folgenden Kapiteln bei jeder Experimentalphase ausführlich dargestellt wird, werden an dieser Stelle nur die schematischen Versuchspläne dargestellt.

Das zusammengefasste Untersuchungsdesign aller drei Phasen ist in Abbildung 3.1.2 schematisch abgebildet. Die einzelnen Versuchsdurchgänge sind zudem in Anhang C (S. 276ff.) einzeln dargestellt.

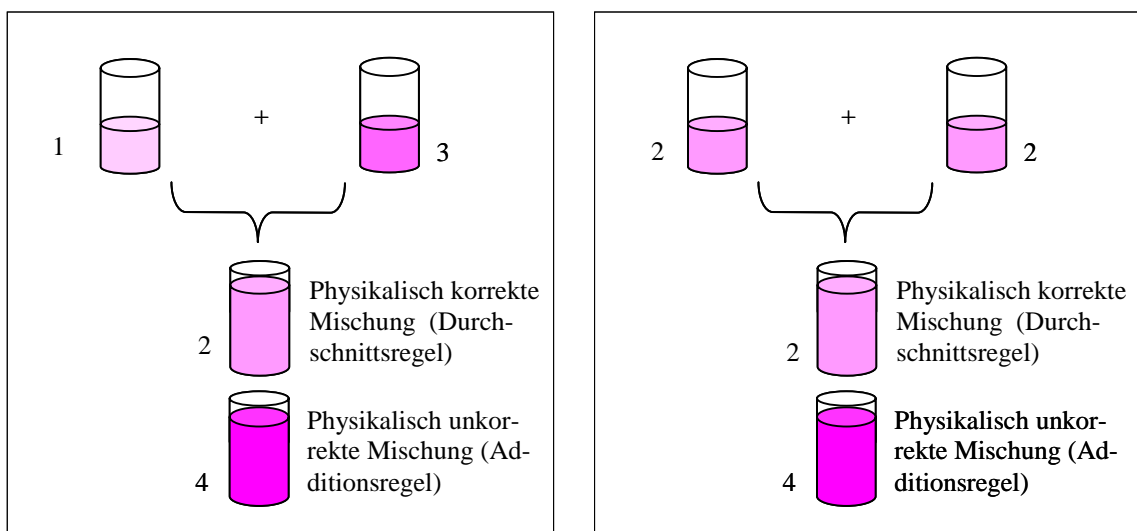


Abbildung 3.1.2. Schematisches zusammengefasstes Untersuchungsdesign der drei Experimentalphasen. Die Zahlen neben den Gläsern entsprechen den verschiedenen Stufen der Farbintensitätsskala.

Allen Phasen gemeinsam waren die quasi-experimentellen Faktoren Alter und Geschlecht. Ausserdem wurden die Versuchsdurchgänge in zwei verschiedenen Reihenfolgen

gen dargeboten, welche pro Altersgruppe gleichmässig den beiden Geschlechtergruppen zugeordnet wurden. Damit konnte ein allfälliger Reihenfolgeeffekt geprüft werden. Zudem wurden die Antwortalternativen bei kategorialen abhängigen Variablen sowie die Instruktionen bei den Skalen ausbalanciert.

Phase 1: Prädiktion der Farbmischungsergebnisse

Unabhängige Variable

- *Faktor A: Farbintensitätskombination (zweistufig)*

In den experimentellen Durchgängen wurden folgende Farbintensitätsstufen kombiniert: Die Intensitätsstufe 2 wurde mit der Intensitätsstufe 2 (gleiche Intensitätskombination) und die Intensitätsstufe 1 mit der Intensitätsstufe 3 (unterschiedliche Intensitätskombination) kombiniert. Beide Kombinationen wurden zweimal dargeboten und das Mischungsurteil mit einer Messwiederholung zweimal pro Kombination erfasst.

Abhängige Variablen

- *Prädiktive Urteile zum Mischungsergebnis: Kategoriale skaliert.*
- *Subjektive Sicherheit des prädiktiven Mischungsurteils: Rating-Skala von 0-10, proportional skaliert (siehe Anhang D, S. 281).*

Phase 2: Magic experiment

Unabhängige Variablen

- *Faktor A: Farbintensitätskombination (zweistufig): Kombinationen 1-3 und 2-2 (siehe Phase 1).*
- *Faktor B: Mischungsart (zweistufig): Physikalisch korrekte und physikalisch unkorrekte Mischungsergebnisse bzw. -arten wurden vorgeführt.*

Die Faktoren wurden durchkombiniert, so dass sich ein 2x2-Untersuchungsdesign mit den Messwiederholungsfaktoren “Farbintensitätskombination” und “Mischungsart” ergab.

Folgende vier Versuchsdurchgänge waren zu beurteilen (vgl. Abbildung 3.1.2 und Anhang C, S. 276ff.):

- *gr* (gleiche Intensitätskombination mit physikalisch korrektem Mischungsresultat),
- *gf* (gleiche Intensitätskombination mit physikalisch unkorrektem Mischungsresultat),
- *ur* (unterschiedliche Intensitätskombination mit physikalisch korrektem Mischungsresultat),
- *uf* (unterschiedliche Intensitätskombination mit physikalisch unkorrektem Mischungsresultat).

Abhängige Variablen

- *Einschätzung des Mischungsresultates auf der Wahrnehmungsebene*: Kategorial skaliert.
- *Einschätzung des Realitätscharakters der Mischung*: Rating-Skala von 0-10, proportionalskaliert (siehe Anhang D, S. 282).
- *Regelsicherheit nach der Mischung*: Rating-Skala von -5 bis +5, intervallskaliert (siehe Anhang D, S. 283).

Phase 3: Kausalitätsverständnis gezauberter Ereignisse

Unabhängige Variablen

- *Faktor A: Farbintensitätskombination (zweistufig)*.
- *Faktor B: Mischungsart (zweistufig)*: Zwei verschiedene Mischungsresultate wurden präsentiert.
- *Faktor C: Machbarkeitszuschreibungen (vierstufig)*: Akteure: Zauberer, Lehrperson, Eltern, Versuchsperson selbst.

Die Faktoren wurden durchkombiniert, so dass sich ein 2x2x4-Untersuchungsdesign mit den Messwiederholungsfaktoren Farbintensitätskombination, Mischungsart und Machbarkeitszuschreibungen ergab.

Abhängige Variable

- *Machbarkeitszuschreibungen*: Rating-Skala von 0 bis 4, intervall- bzw. ordinalskaliert.

3.2 Prädiktion der Farbmischungsresultate

3.2.1 Fragestellung

Die bestehende Befundlage zu den kindlichen Voraussagefähigkeiten eines Mischungsresultates nach dem Zusammenleeren von Flüssigkeiten mit intensiven Grössen ist kontrovers. Die früheren Studien von Strauss und Mitarbeitern (v. a. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982) mit Mischungen der Kombinationen gleich intensiver Grössen sagen einen u-förmigen Entwicklungsverlauf der korrekten Mischungsregel voraus. Das Experiment von Jäger (1996, Experiment 1) hingegen zeigte nur bei der Mischung unterschiedlicher Intensitätskombinationen einen u-förmigen Verlauf. In einem weiteren Experiment mit unterschiedlichen Intensitätskombinationen von Jäger (1996, Experiment 2 bzw. Jäger & Wilkening, 2001, Experiment 1) stieg der Anteil korrekter Voraussagen allerdings linear an. Die Sechsjährigen urteilten meistens inkonsistent, die Acht- und Zehnjährigen addierten hauptsächlich und die Zwölfjährigen schliesslich wendeten am meisten die Durchschnittsregel an. Die Befundlage bei der Mischung gleicher Intensitätskombinationen (Jäger, 1996, Experiment 1) zeigte bereits bei den jüngsten Kindern eine hohe Anzahl korrekter Lösungen. Allerdings war auch hier der Entwicklungsanstieg nicht linear. Jäger führte die verbesserte Voraussagefähigkeit bei der Mischung gleicher Intensitäten auf die Darbietung einer direkt wahrnehmbaren intensiven Grösse in Form von Farbintensitäten (Farbhelligkeiten) zurück. Auf der Basis dieser Argumentation wären bei der vorliegenden Untersuchung ähnliche Verläufe wie bei Jäger zu erwarten. Allerdings unterschieden sich in den Vorläuferstudien auch die abhängigen Variablen. Während die Versuchspersonen bei Jäger die Antworten auf einer 7-teiligen Farbskala einschätzen mussten, wurden die Antworten bei Strauss (v. a. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982) wie auch in der vorliegenden Untersuchung in Phase 1 mit der Paarvergleichsmethode erfasst.

Auf der Basis dieser Grundlage können folgende zentralen Fragestellungen formuliert werden:

- Werden bei beiden Intensitätskombinationen vorwiegend Antworten gemäss der Additions- und Durchschnittsregel gegeben?

- Urteilen die Kinder bei der Mischung gleicher Intensitäten häufiger korrekt als bei der Mischung unterschiedlicher Intensitäten?
- Wie sieht der Entwicklungsverlauf der verwendeten Regeln aus? Steigen die Häufigkeiten korrekten Durchschnittsbildungen mit zunehmendem Alter stetig an oder verlaufen sie u-förmig?

Hinsichtlich der Einschätzungen der subjektiven Sicherheit des prädiktiven Mischungsurteils wird erwartet, dass jüngere Kinder weniger konsistent urteilen, da die eben erworbenen introspektiven Fähigkeiten vermutlich noch wenig präzise Einschätzungen ermöglichen. Hinter der Erfassung der subjektiven Regelsicherheit stecken aber vielmehr folgende Fragen: Steigt die Gewissheit der korrekten Regel mit zunehmendem Alter an? Unterscheiden sich die verschiedenen Regelanwender in ihrer subjektiven Sicherheit? Zudem war es durch die Erfassung eines individuellen Sicherheitsniveaus möglich, einen allfälligen Einfluss des Grundniveaus auf die Sicherheitseinschätzung nach Beobachtung der Mischungsvorgänge durch eine Kovarianzanalyse zu eliminieren.

3.2.2 Methode

Versuchsplan

Da die Untersuchung der Prädiktion mit der Paarvergleichsmethode vor der Mischung der Flüssigkeiten durchgeführt wurde, bezieht sich das Untersuchungsdesign nur auf die Variation in der Intensitätskombinationen. Pro Intensitätskombination (gleiche und unterschiedliche) erfolgten zwei Versuchsdurchgänge, d. h. die abhängigen Variablen zur Prädiktion mit der Paarvergleichsmethode wurden je zweimal unter der gleichen Bedingung erhoben.

Unabhängige Variablen

- Faktor A (Messwiederholungsfaktor; zweistufig): Farbintensitätskombination (Stimulusdarbietung siehe Anhang C, S. 276ff.)
- Faktor B (Zwischensubjektfaktor; fünfstufig): Alter
- Faktor C (Zwischensubjektfaktor; zweistufig): Geschlecht

Abhängige Variablen

- *Erfassung der prädiktiven Mischungsurteile:* Wird es heller, dunkler oder bleibt es gleich? Antwortalternativen wurden in der Reihenfolge variiert; nominal bzw. ordinal.
- *Subjektive Sicherheit des prädiktiven Mischungsurteils:* Wie sicher bist du, dass es ... wird? Skala von 0-10 (siehe Anhang D, S. 281), Erklärung der Skala mit variiertem Reihenfolge, intervallskaliert.

3.2.3 Auswertung

Es wurde sowohl eine Auswertung der prädiktiven Mischungsurteile als auch eine Auswertung der subjektiven Sicherheiten dieser Mischungsurteile vorgenommen.

Prädiktion der Farbmischungsurteile mit der Paarvergleichsmethode

Zuweisung der Urteile zu Mischungsregeln

Die Urteile der Versuchspersonen aus der Paarvergleichsmethode wurden einzelnen Mischungsregeln zugeordnet.

Da aufgrund von Jägers Ergebnissen v. a. mit der Verwendung der Durchschnitts- und Additionsregel (bei Jäger, 1996: Intensivierungsregel) zu rechnen war und in der vorliegenden Untersuchung nur die Regelzuteilung der einzelnen Versuchspersonen wesentlich war, wurden die Mischungsregeln durch die Wahl vorgegebener Kategorien erfasst. Die Verwendung der Begriffe „...regel“ suggeriert demnach einen Zusammenhang zwischen der Wahl einer Kategorie und der Verknüpfung der Informationen nach einer bestimmten Regel, der nicht aus der vorliegenden Untersuchung hervorgeht.

Die je drei möglichen Antworten pro Intensitätskombination wurden folgenden Regeln zugewiesen:

Gleiche Intensitätskombination

- Mischungsergebnis blieb gleich: Durchschnittsregel
- Mischungsergebnis wurde dunkler: Additionsregel

- Mischungsergebnis wurde heller: Subtraktionsregel

Unterschiedliche Intensitätskombination

- Mischungsergebnis wurde heller als das dunklere der beiden: Durchschnittsregel
- Mischungsergebnis wurde dunkler als das dunklere der beiden: Additionsregel
- Mischungsergebnis blieb gleich wie das dunklere der beiden: Zentrierungsregel

Kategorisierte, nominalskalierte Mischungsurteile

Wurde die gleiche Regel zweimal innerhalb derselben Intensitätskombinationsbedingung verfolgt, so wurden die Mischungsurteile als konsistente Regel kategorisiert. Alle anderen gemischten Urteile wurden der inkonsistenten Regel zugeordnet. Die Ratewahrscheinlichkeit für eine konsistente Regel entsprach einem Neuntel, diejenige für die inkonsistente Mischungsregel sechs Neunteln oder zwei Dritteln. Mit dem χ^2 -Einzeltest soll ausgeschlossen werden, dass die beobachteten Verteilungen durch Raten entstanden waren.

Die Verteilungen der konsistent angewendeten Regeln über die verschiedenen Altersgruppen wurden mit dem χ^2 -Test nach Pearson auf Unterschiede geprüft. Allerdings könnte die Voraussetzung für diesen Test, dass nur maximal 20 Prozent der Zellen eine erwartete Häufigkeit < 5 haben dürfen, nicht erfüllt gewesen sein. In diesem Falle wurde die Berechnung nur mit denjenigen Regeln durchgeführt, welche genügend besetzt waren, damit das Resultat interpretierbar war. Zur Interpretation der Ergebnisse eines χ^2 -Tests tragen wesentlich die standardisierten Residuen in der Kreuztabelle bei, da die Felder mit hohen standardisierten Residuen einen wesentlichen Beitrag zum χ^2 -Wert und damit zu einem allfällig signifikanten Ergebnis liefern. Üblicherweise deutet ein standardisiertes Residuum > 2 auf ein signifikantes Resultat.

Weiter wurde mit dem χ^2 -Test nach Pearson geprüft, ob die Verteilungen der Geschlechtszugehörigkeit bzw. Reihenfolge der Experimentaldurchgänge unabhängig von der Verteilungen der prädiktiven Mischungsregeln waren. Zudem wurden die Verteilungen korrekter und unkorrekter Regeln bei den verschiedenen Intensitätskombinationen mit dem Vierfelder- χ^2 -Test nach Pearson auf Unabhängigkeit geprüft.

Kategorisierte ordinalskalierte Mischungsurteile

Pro Regelkategorie (Durchschnitts- und Additionsregel sowie die Restkategorie) waren drei Antworten möglich: nie genannt (0), einmal genannt (1) oder zweimal genannt (2). Die erfassten Daten lagen innerhalb der einzelnen Regelkategorien also auch ordinalskaliert (0,1 oder 2) vor, was weitere statistische Berechnungen zuliess. So konnte mit dem Kruskal-Wallis-Test für mehrere unabhängige Stichproben überprüft werden, ob sich die Ausprägungen der einzelnen Regeln im Entwicklungsverlauf unterschieden hatten. Mit dem Friedman wurde untersucht, ob sich die Ausprägungen der verschiedenen Regeln signifikant unterschieden hatten. Mit dem Wilcoxon-Test für zwei abhängige Stichproben schliesslich wurden die Ausprägungen der einzelnen Regeln der gleichen Intensitätskombination mit denjenigen bei der unterschiedlichen Intensitätskombination verglichen.

Subjektive Sicherheit des prädiktiven Mischungsurteils

Konsistenzanalyse

Da die subjektiven Sicherheitseinschätzungen der prädiktiven Mischungsurteile innerhalb einer Intensitätskombination zweimal erhoben wurden, wurden deren Konsistenzen mit der Korrelationsanalyse nach Pearson berechnet. Da erwartet wurde, dass die Messwiederholungsdaten in linear positivem Zusammenhang stehen, genügte eine einseitige Hypothesenprüfung. Nebst den Korrelationen innerhalb der beiden Intensitätskombinationen war auch der Zusammenhang der subjektiven Regelsicherheiten zwischen den beiden Intensitätskombinationen von Interesse.

Varianzanalysen

Zur Überprüfung eines Reihenfolge-Effektes (die Daten wurden mit zwei verschiedenen Reihenfolgen der Versuchsdurchgänge erhoben) auf die subjektive Regelsicherheit wurde eine Varianzanalyse mit den Messwiederholungsfaktoren Intensitätskombination und Messzeitpunkt sowie den Zwischensubjektfaktoren Reihenfolge und Alter durchgeführt.

Zentral war aber die Durchführung einer dreifaktoriellen Varianzanalyse für die mittlere subjektive Regelsicherheit mit dem Messwiederholungsfaktor Farbintensitätskombina-

tion sowie den Zwischensubjektfaktoren Alter und Geschlecht. Auftretende Wechselwirkungen wurden weiter berechnet und analysiert.

Um zu prüfen, ob sich die zu Regeln kategorisierten prädiktiven Mischungsurteile auf das zugehörige subjektive Sicherheitsurteil auswirkten, wurden weitere Varianzanalysen berechnet. Da sich die Einschätzungen nur innerhalb je einer Intensitätskombination zu konsistenten Regelanwendungen zuordnen liessen, mussten die Varianzanalysen für die einzelnen Intensitätskombinationen getrennt durchgeführt werden. Es wurde pro Intensitätskombination je eine dreifaktorielle univariate Varianzanalyse mit den Zwischensubjektfaktoren prädiktive Regelanwendung, Alter und Geschlecht berechnet. Bei auftretenden Wechselwirkungen zwischen dem Faktor Alter und einem der anderen Faktoren wurden für die einzelnen Altersstufen getrennte Varianzanalysen berechnet.

3.2.4 Ergebnisse

Prädiktive Mischungsregeln

Die Häufigkeiten und Prozente der zu konsistenten Regeln kategorisierten Mischungsurteile sind für beide Intensitätskombinationen in Tabelle 3.2.1 dargestellt.

Tabelle 3.2.1

Häufigkeiten und Prozente der konsistent angewendeten Mischungsregeln in Abhängigkeit von der Altersgruppe der Versuchspersonen und der Intensitätskombination der Ursprungsflüssigkeiten (gleiche, unterschiedliche)

Altersgruppe	Mischungsregel							
	Inkonsistenz		Durchschnitt		Addition		Restkategorie	
	gleiche	unterschiedl.	gleiche	unterschiedl.	gleiche	unterschiedl.	gleiche	unterschiedl.
Sechsjährige (n=30)	10 (33 %)	12 (40 %)	11 (37 %)	9 (30 %)	7 (23 %)	8 (27 %)	2 (7 %)	1 (3 %)
Achtjährige (n=30)	9 (30 %)	8 (27 %)	15 (50 %)	15 (50 %)	6 (20 %)	7 (23 %)	0 (0 %)	0 (0 %)
Zehnjährige (n=30)	12 (40 %)	3 (10 %)	10 (33 %)	23 (77 %)	6 (20 %)	4 (13 %)	2 (7 %)	0 (0 %)
Zwölfjährige (n=29)	3 (10 %)	4 (13 %)	19 (63 %)	23 (77 %)	8 (27 %)	3 (10 %)	0 (0 %)	0 (0 %)
Erwachsene	0	0	20	18	0	2	0	0

(n=20)	(0 %)	(0 %)	(100 %)	(90 %)	(0 %)	(10 %)	(0 %)	(0 %)
Alle Vpn	34	27	75	88	27	24	4	1
(n=140)	(24 %)	(19 %)	(54 %)	(63 %)	(19 %)	(17 %)	(3 %)	(1 %)

Anmerkung. Restkategorie bei der gleichen Intensitätskombination entspricht der Subtraktionsregel und bei der unterschiedlichen Intensitätskombination der Zentrierungsregel.

Zur Überprüfung, ob die konsistenten Regelanwendungen zufällig durch Raten entstanden sein konnten, wurden die Häufigkeiten der konsistenten Anwendung der Durchschnitts- und Additionsregeln sowie der Restkategorie (Ratewahrscheinlichkeit von je einem Neuntel) gegenüber der inkonsistenten Regelanwendung (Ratewahrscheinlichkeit von sechs Neuntel) mit dem χ^2 -Einzeltest überprüft. Sowohl bei der gleichen als auch der unterschiedlichen Intensitätskombination zeigten sich aber höchst signifikante Abweichungen der konsistenten bzw. inkonsistenten Regelanwendung von der Ratewahrscheinlichkeit ($\chi^2(3, N = 140) = 281.89, p < .001$ bzw. $\chi^2(3, N = 140) = 402.73, p < .001$). Die Überprüfung bei den jüngeren Kindern (Sechs- und Achtjährige) wies ebenfalls bei beiden Farbintensitätskombinationen höchst signifikante Unterschiede zwischen erwarteten und beobachteten Verteilungen auf ($\chi^2(3, N = 60) = 76.37, p < .001$ bzw. $\chi^2(3, N = 60) = 70.3, p < .001$).

Regelanwendung bei der gleichen Intensitätskombination

Konsistente Regelanwendung

In Abbildung 3.2.1 sind die prozentualen Anteile der Versuchspersonen in allen Altersstufen, die bei der Mischung gleicher Farbintensitäten die einzelnen Mischungsregeln befolgten, dargestellt.

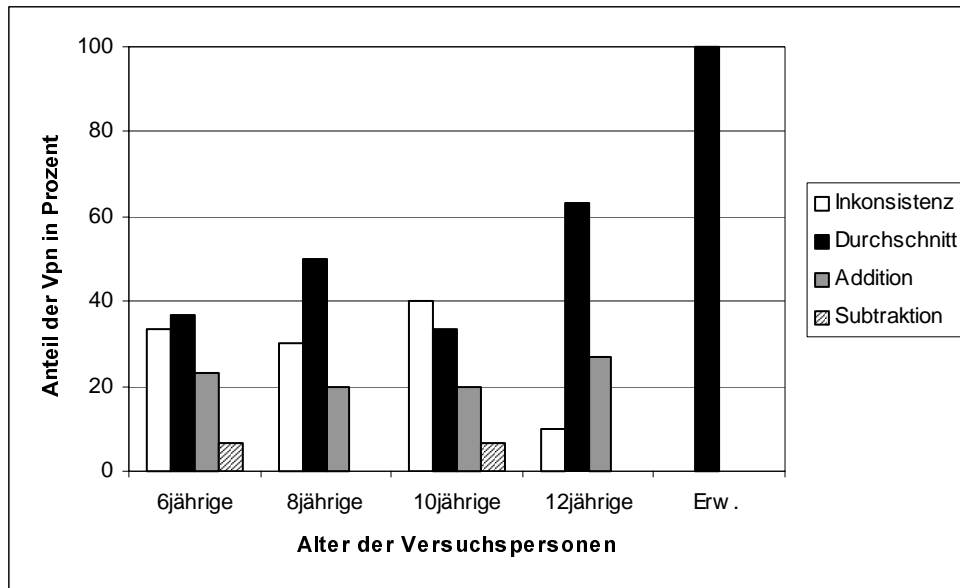


Abbildung 3.2.1. Häufigkeiten der konsistenten und inkonsistenten Regelanwendungen bei der gleichen Intensitätskombination.

Wie aus Tabelle 3.2.1 und Abbildung 3.2.1 ersichtlich wird, wendeten nur gerade je zwei Versuchspersonen aus den Altersgruppen der Sechs- und Zehnjährigen die Subtraktionsregel konsistent an. 19 Prozent aller Versuchspersonen (27 Kinder) sagten konsistent ein Mischungsresultat voraus, das dunkler werden würde. Die Addierer waren auf die Altersgruppen der Kinder in etwa gleich verteilt. Etwas häufiger als addiert, wurde inkonsistent geantwortet (24 Prozent der Versuchspersonen bzw. 34 Kinder). Hier fiel auf, dass nur noch wenige der Zwölfjährigen zu den inkonsistenten Regelanwendern zählten und sich unter den Zehnjährigen die meisten inkonsistenten befanden. Reziprok zu den Häufigkeiten der inkonsistenten Anwender verläuft die Kurve bei den Anwendern der Durchschnittsregel. Nur wurde korrekte Voraussage weit häufiger gemacht (54 Prozent aller Versuchspersonen bzw. 75 Versuchspersonen). Mit Ausnahme eines Einbruches bei den Zehnjährigen nahm die Anzahl der Durchschnittsbilder mit dem Alter stetig zu und erreichte bei den Erwachsenen eine Häufigkeit von 100 Prozent.

Die genaue Analyse der Zusammensetzung verschiedener Regeln bei den inkonsistenten Regelanwendern zeigte, dass von insgesamt 34 Inkonsistenten die Mehrheit, nämlich 20 (59 % aller Inkonsistenten) einmal die Durchschnittsregel und einmal die Additionsregel angewendet hatte. Die restlichen Inkonsistenten verteilten sich auf die Anwendung von Durchschnitts- und Subtraktionsregel sowie Additions- und Subtraktionsregel (je 7 Versuchspersonen, d. h. 21 % aller Inkonsistenten).

Die statistische Überprüfung der Unabhängigkeit der Verteilungen der konsistenten Regelanwendungen bezüglich des Alters wurde aufgrund der zu geringen Zellenbesetzung der erwarteten Häufigkeiten bei der Subtraktionsregel ohne diese Kategorie berechnet. Die Berechnung des χ^2 -Tests nach Pearson ergab bei den gleichen Intensitäten einen höchst signifikanten χ^2 -Wert ($\chi^2(8, N = 136) = 36.78; p < .001$). Die Kreuztabelle zeigte, dass die Erwachsenen (20 Erwachsene bzw. 100 %; standardisiertes Residuum = 2.7) besonders häufig richtig und die Zehnjährigen (12 von 28 Kindern bzw. 43 %; standardisiertes Residuum = 1.9) besonders häufig inkonsistent urteilten. Die Überprüfung, ob die beiden Verteilungen auf die verschiedenen Mischungsregeln und die Geschlechtszugehörigkeit bzw. Reihenfolge der Experimentaldurchgänge voneinander abhängig waren, zeigte keinen Zusammenhang der Verteilungen bei beiden Variablen.

Mittlere Regelanwendung

Die pro Intensitätskombination zweimal erhobenen Mischungsurteile konnten auch hinsichtlich ihrer mittleren Anwendung pro Regel (null, einmal oder zweimal angewendete Regel) ausgewertet werden. Abbildung 3.2.2 zeigt das Liniendiagramm für die mittleren Anwendungen der verschiedenen Regeln bei der gleichen Intensitätskombination.

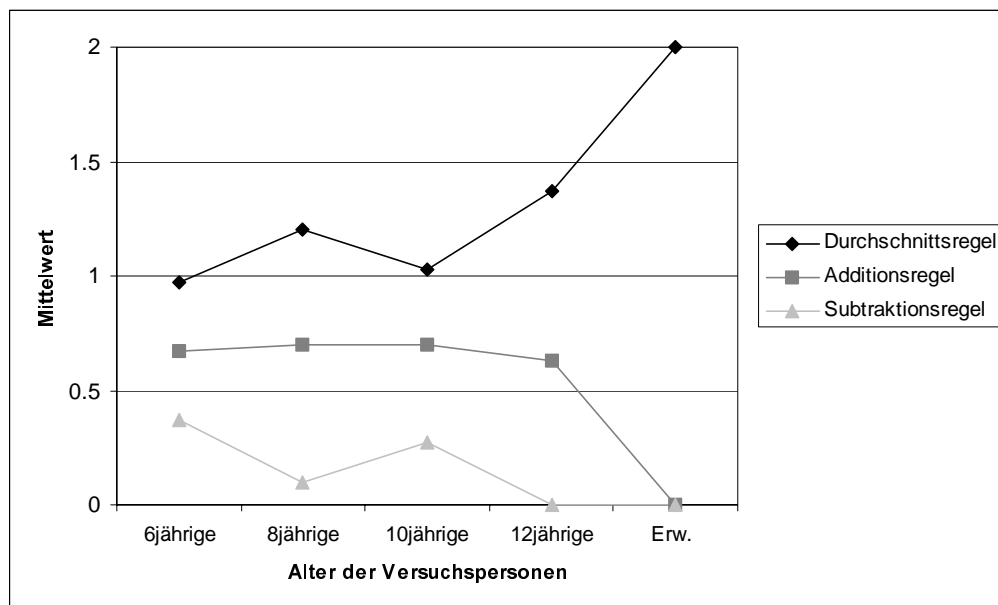


Abbildung 3.2.2. Mittelwerte der angewendeten Regeln (Durchschnitts-, Additions- und Subtraktionsregel) bei der gleichen Intensitätskombination in Abhängigkeit von der Altersgruppe (Max.=2).

Mit dem Kruskal-Wallis-Test für unabhängige Stichproben wurden die mittleren Ränge der beiden häufigsten Regeln, der Durchschnitts- und Additionsregel, über die verschiedenen Altersgruppen miteinander verglichen. Es zeigte sich, dass die richtige Durch-

schnittsregel im Altersverlauf höchst signifikant verschieden angewendet wurde, $\chi^2(4, N = 140) = 22.9; p < .001$. Wie obige Abbildung zeigt, verlief der Entwicklungsverlauf der mittleren richtigen Regelanwendung bei den gleichen Intensitäten aber unstetig, da die Zehnjährigen tiefer waren als die Achtjährigen. Auch die Additionsregel wurde über das Alter hinweg sehr signifikant verschieden häufig angewendet, $\chi^2(4, N = 140) = 14.26; p < .01$. Abbildung 3.2.2 macht deutlich, dass dieses Ergebnis nur auf den Unterschied zwischen den Kindern und den Erwachsenen zurückzuführen ist.

Die Berechnung mit dem Friedman-Test für mehrere abhängige Variablen (die drei verschiedenen Regeln) zeigte, dass sich die Anwendungen der Durchschnitts-, Additions- und Subtraktionsregeln bei den gleichen Intensitäten sehr deutlich unterschieden, $\chi^2(2, N = 140) = 76.11; p < .001$. Weitere Berechnungen mit dem Wilcoxon-Test auf der Datenbasis der Kinder zeigten, dass sowohl die Durchschnitts- und die Additionsregel ($z(120) = 3.09, p < .01$; 2-seitig) als auch die Additions- und die Subtraktionsregel ($z(120) = 4.68, p < .001$; 2-seitig) höchst signifikant verschieden angewendet wurden

Regelanwendung bei der unterschiedlichen Intensitätskombination

Konsistente Regelanwendung

Die Abbildung 3.2.3 zeigt die prozentualen Anteile der Versuchspersonen aus den fünf Altersstufen der Anwendung der verschiedenen Mischungsregeln für die Aufgabe der Mischung unterschiedlicher Farbintensitäten.

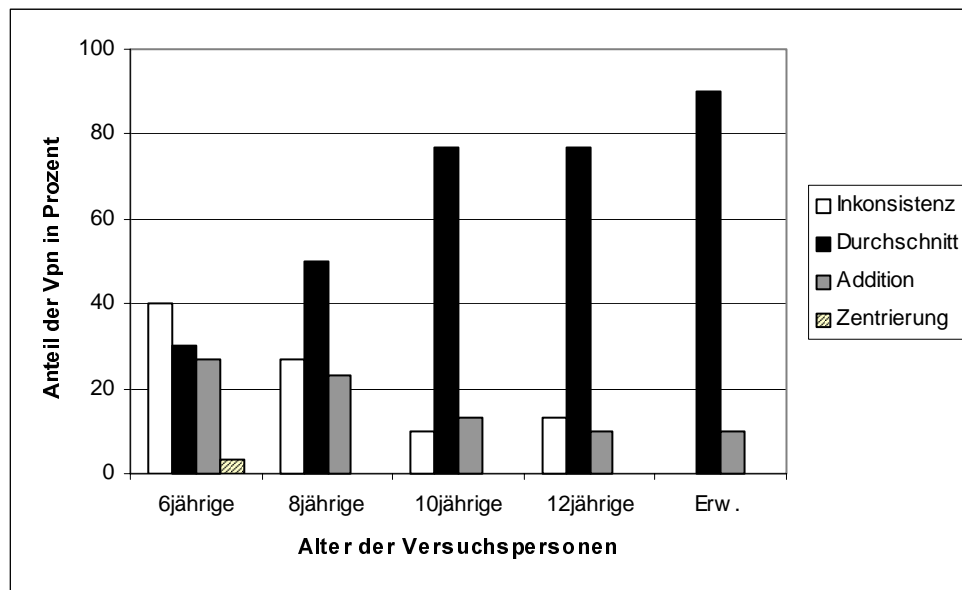


Abbildung 3.2.3. Häufigkeiten der inkonsistenten und inkonsistenten Regelanwendungen bei der unterschiedlichen Intensitätskombination.

Tabelle 3.2.1 und Abbildung 3.2.3 zeigen, dass die Zentrierungsregel nur gerade von einer Versuchsperson aus der Altersgruppe der Sechsjährigen konsistent angewendet wurde. Weit am grössten ist der Anteil der Versuchspersonen, welche die Durchschnittsregel anwendeten (63 % aller Versuchspersonen bzw. 88 Versuchspersonen). Während bei den Achtjährigen schon 50 % die richtige Durchschnittsregel konsistent anwendeten, stieg die Häufigkeit der richtigen Regel mit dem Alter noch deutlich an (bereits 77 % der Zehn- und Zwölfjährigen) und erreichte bei den Erwachsenen 90 % (eine erwachsene Person wendete die Additionsregel an). Inkonsistente Regeln und Additionsregeln wurden deutlich weniger häufig angewendet (19 bzw. 17 % aller Versuchspersonen). Beide Kurven nehmen mit zunehmendem Alter ab und verlaufen reziprok zur Kurve der konsistenten Durchschnittsbildung.

Die genaue Analyse der Zusammensetzung verschiedener Regeln bei den inkonsistenten Regelanwendern zeigte, dass von insgesamt 26 Inkonsistenten die Mehrheit, nämlich 16 (62 % aller Inkonsistenten) einmal die Durchschnittsregel und einmal die Additionsregel angewendet hatte. Die restlichen Inkonsistenten verteilten sich auf die Anwendung von Durchschnitts- und Subtraktionsregel sowie Additions- und Subtraktionsregel (je 5 Versuchspersonen, d. h. 19 % aller Inkonsistenten).

Die statistische Überprüfung der Unabhängigkeit der Verteilungen der konsistenten Regelanwendungen bezüglich des Alters wurde aufgrund der zu geringen Zellenbesetzung der erwarteten Häufigkeiten bei der Zentrierungsregel ohne diese Kategorie berechnet. Die Berechnung des χ^2 -Tests nach Pearson ergab bei den unterschiedlichen Intensitäten einen höchst signifikanten χ^2 -Wert, $\chi^2(8, N = 139) = 30.77$; $p = .001$. Die Kreuztabelle wies darauf hin, dass die Erwachsenen (keine einzige Person; standardisiertes Residuum = -2.0) besonders selten inkonsistent sowie die Sechsjährigen besonders häufig inkonsistent (12 von 29 Kindern bzw. 41 %; standardisiertes Residuum = 2.7) und besonders selten richtig (9 von 29 Kindern bzw. 31 %; standardisiertes Residuum = -2.2) urteilten. Die Überprüfung, ob die beiden Verteilungen auf die verschiedenen Mischungsregeln und die Geschlechtszugehörigkeit bzw. Reihenfolge der Experimentaldurchgänge voneinander abhängig waren, zeigte keinen Zusammenhang der Verteilungen bei beiden Variablen.

Mittlere Regelanwendung

Die pro Intensitätskombination zweimal erhobenen Mischungsurteile konnten auch hinsichtlich ihrer mittleren Anwendung pro Regel (null, einmal oder zweimal angewendete Regel) ausgewertet werden. Abbildung 3.2.4 zeigt das Liniendiagramm für die mittleren Anwendungen der verschiedenen Regeln bei der unterschiedlichen Intensitätskombination.

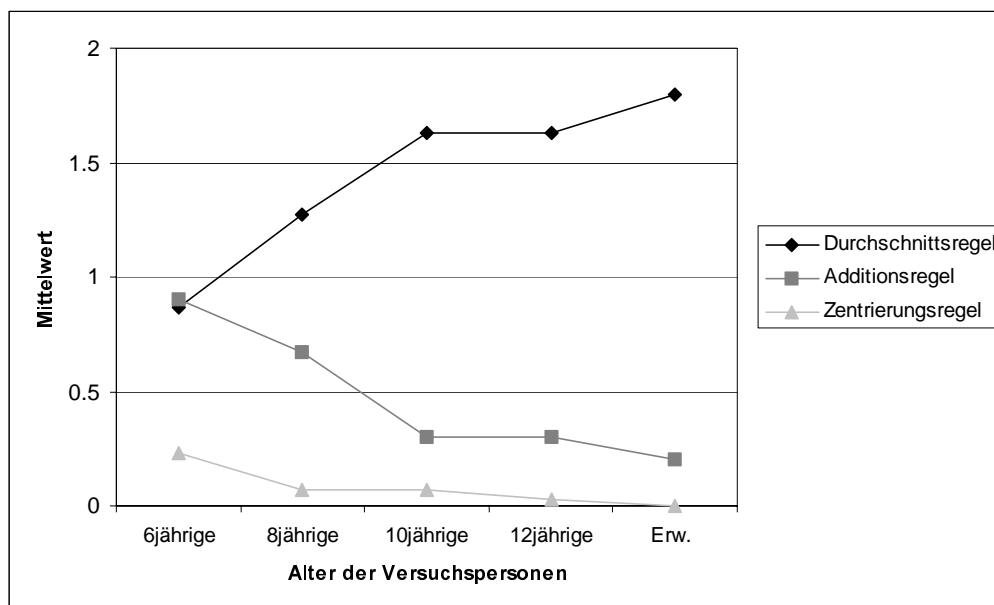


Abbildung 3.2.4. Mittelwerte der angewendeten Regeln (Durchschnitts-, Additions- und Zentrierungsregel) bei der unterschiedlichen Intensitätskombination (Max.=2).

Vergleicht man obige Abbildung mit den in Abbildung 3.2.2 dargestellten Mittelwerten der angewendeten Regeln bei der gleichen Intensitätskombination so fällt auf, dass bei der unterschiedlichen Intensitätskombination sich die Anwendungen von Durchschnitts- und Additionsregeln bereits bei den Kindern veränderten: Während die Additionsregel bei der gleichen Intensitätskombination bei den Kindern konstant blieb, nahm sie bei der unterschiedlichen Intensitätskombination stetig ab.

Mit dem Kruskal-Wallis-Test für unabhängige Stichproben wurden die mittleren Ränge der beiden häufigsten Regeln, der Durchschnitts- und Additionsregel, über die verschiedenen Altersgruppen miteinander verglichen. Es zeigte sich, dass die richtige Durchschnittsregel im Altersverlauf höchst signifikant verschieden angewendet wurde, $\chi^2(4, N = 140) = 25.3; p < .001$. Wie obige Abbildung zeigt, verlief der Entwicklungsverlauf der mittleren richtigen Regelanwendung bei den gleichen Intensitäten mit zunehmendem Alter linear ansteigend. Auch die Additionsregel wurde über das Alter hinweg sehr signifikant verschieden häufig angewendet, $\chi^2(4, N = 140) = 21; p < .001$. Abbildung 3.2.4 zeigt, dass die mittlere Anwendung der Additionsregel fast komplementär zu derjenigen der Durchschnittsregel verlief.

Die Berechnung mit dem Friedman-Test für mehrere abhängige Variablen (die drei verschiedenen Regeln) zeigte, dass sich die Anwendungen der Durchschnitts-, Additions- und Zentrierungsregeln bei den unterschiedlichen Intensitäten sehr deutlich unterscheiden, $\chi^2(2, N = 140) = 108.44; p < .001$. Weitere Berechnungen mit dem Wilcoxon-Test auf der Datenbasis der Kinder zeigten, dass sowohl die Durchschnitts- und die Additionsregel ($z(120) = 5.01, p < .001$; 2-seitig) als auch die Additions- und die Subtraktionsregel ($z(120) = 4.9, p < .001$; 2-seitig) höchst signifikant verschieden angewendet wurden.

Vergleich der beiden Intensitätskombinationen

Unkorrekte und konsistent korrekte Regelanwendung

In Abbildung 3.2.5 sind die prozentualen Anteile der Versuchspersonen in allen Altersstufen, die bei den beiden Intensitätskombinationen konsistent die richtigen Mischungsregeln befolgten, dargestellt.

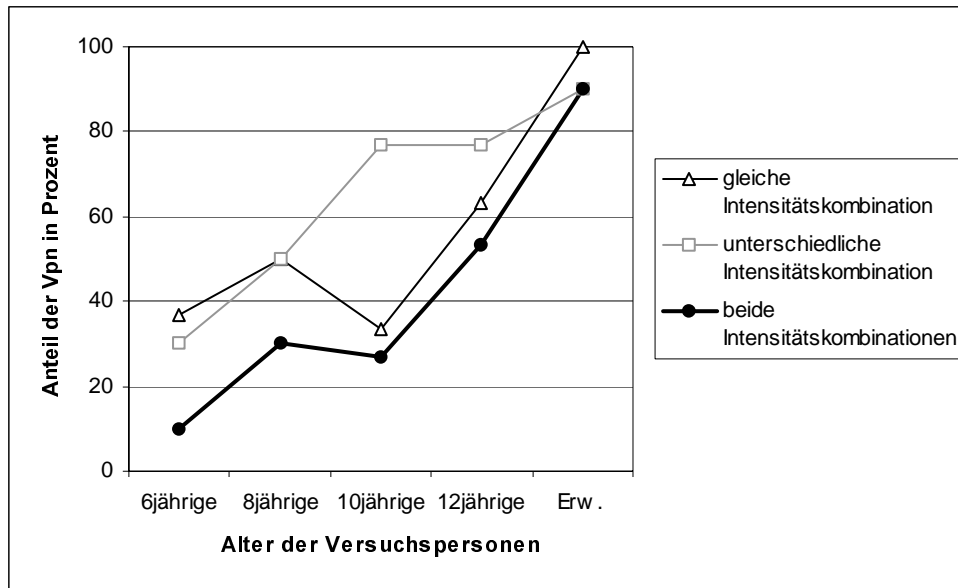


Abbildung 3.2.5. Prozentualer Anteil der konsistent richtigen Regelanwendungen bei der gleichen und der unterschiedlichen Intensitätskombination sowie bei beiden Intensitätskombinationen pro Altersgruppe.

Die Abbildung 3.2.5 zeigt die Unterschiede bei den Verlaufsformen der konsistent korrekten Regelanwendungen zwischen den beiden Intensitätskombinationen, die insbesondere durch den geringen Anteil korrekter Urteile bei der Altersgruppe der Zehnjährigen zu Stande kamen. Der Anstieg der angewendeten Durchschnittsregel bei beiden Intensitätskombinationen verlief mit Ausnahme der Zehnjährigen mit zunehmendem Alter linear.

Die Überprüfung, ob sich die Häufigkeiten korrekter und unkorrekter Regeln bei den beiden Intensitätskombinationen (Kategorisierung in unkorrekte und konsistent korrekte Regeln) unterschieden, mit dem Vierfelder- χ^2 -Test zeigte einen signifikanten Unterschied, $\chi^2(1, 140) = 5.78, p < .05$. Weitere Vierfelder- χ^2 -Tests bei den Altersgruppen der Zehn- und Zwölfjährigen ergaben nur bei den Zehnjährigen eine höchst signifikante Abhängigkeit der korrekten konsistenten Mischungsurteile von der Intensitätskombination, $\chi^2(1, 30) = 11.38, p < .001$.

Zudem wurde noch analysiert, wie viele der Versuchspersonen bei beiden Intensitätskombinationen unkorrekt oder konsistent korrekt geantwortet hatten. Die Häufigkeiten der Versuchspersonen, welche sowohl bei den gleichen als auch den unterschiedlichen Intensitätskombinationen falsche Mischungsergebnisse voraussagten, nahmen mit zunehmendem Alter stetig ab. 42 % von allen doppelt falschen Anwendern befanden sich un-

ter den Sechsjährigen (43 % der Sechsjährigen). Die Achtjährigen machten noch 29 %, die Zehnjährigen noch 16 % und die Zwölfjährigen noch 13 % aller Anwender aus, welche sowohl bei der gleichen als auch der unterschiedlichen Intensitätskombination falsch urteilten. Der Verlauf der Häufigkeiten der konsistent richtigen Antworten hingegen nahm mit zunehmendem Alter ansteigend zu, allerdings antworteten die Zehnjährigen weniger häufig bei allen vier Versuchsdurchgängen richtig, als die Achtjährigen. Bei den Zwölfjährigen urteilten 53 % der Altersgruppe bei beiden Intensitätskombinationen richtig, bei den Erwachsenen waren es 90 %, da zwei Personen bei der gleichen Intensitätskombination falsch antworteten. Der χ^2 -Test nach Pearson mit den drei Kategorien „zweimal richtig“, „einmal richtig/falsch“ und „zweimal falsch“ sowie den Altersgruppen zeigte einen – aufgrund der sehr deutlichen Altersunterschiede der doppelt richtigen bzw. falschen Anwendungen wenig überraschenden – höchst signifikanten Unterschied, $\chi^2(8, N = 140) = 43,6; p < .001$.

Mittlere Regelanwendung

Der Vergleich der Anwendung der Durchschnittsregel über beide Intensitätskombinationen zeigte nur bei der Stichprobe der Kinder bedeutsame Unterschiede (Wilcoxon-Test: $z(120) = 2.01; p > .05$, 2-seitig). Die beiden Vergleiche der Additionsregel bzw. der Restkategorien über beide Intensitätskombinationen fielen nicht signifikant bzw. mit einem Trend aus.

Subjektive Sicherheit der Mischungsurteile vor dem Mischungsvorgang

Konsistenzen

Die Durchführung von Korrelationsanalysen nach Pearson ergab bei beiden Intensitätskombinationen höchst signifikante Konsistenzen: Die beiden Sicherheitsurteile korrelierten bei der gleichen Intensitätskombination leicht höher ($r = .76, p < .001$) als bei der unterschiedlichen Intensitätskombination ($r = .61, p < .001$). Der Zusammenhang zwischen den beiden Intensitätskombinationen war zwar weniger eng, aber immer noch sehr signifikant ($r = .54, p < .001$).

Die deskriptiven Daten zeigten, dass die Streuungen der subjektiven Sicherheitseinschätzungen der prädiktiven Mischungsurteile mit abnehmendem Alter stetig grösser wurden (siehe Tabelle 3.2.2). Um zu überprüfen, ob diese z. T. recht grossen Streuun-

gen auch mit einer geringeren Konsistenz einher gingen, wurden die Konsistenzen der Sicherheitsurteile zudem noch pro Altersgruppen berechnet und in Tabelle 3.2.2 dargestellt.

Tabelle 3.2.2

Korrelationskoeffizienten und Signifikanztest (einseitige Hypothesenprüfung) sowie Standardabweichungen der Mittelwerte der beiden Messwiederholungen der subjektiven Regelsicherheit pro Intensitätskombination der Altersgruppe der Versuchspersonen

Altersgruppe	gleiche Intensitätskombination			unterschiedliche Intensitätskombination		
	s (gr)	s (gf)	Korrelation	s (ur)	s (uf)	Korrelation
Sechsjährige (n=30)	2.38	2.72	.71***	3.24	3.11	.45*
Achtjährige (n=30)	2.44	2.46	.77***	2.65	2.93	.53**
Zehnjährige (n=30)	1.88	2.34	.58**	2.19	2.19	.81***
Zwölfjährige (n=29)	1.63	1.80	.76***	1.63	2.12	.38*
Erwachsene (n=20)	.49	.49	1.00 ⁻	1.34	1.34	1.00 ⁻

Anmerkung. Bei der perfekten Korrelation, wie sie sich bei den Erwachsenen zeigte, konnte kein Signifikanztest berechnet werden.

Der obigen Tabelle ist zu entnehmen, dass die hohen Standardabweichungen nicht auf mangelnde Konsistenz bei den jüngeren Kindern zurückzuführen waren, wie sich besonders deutlich an den starken Korrelationen bei den jüngeren Kindern in der Bedingung der gleichen Intensitätskombination illustrieren lässt.

Es zeigten sich ausserdem auch Unterschiede in den Konsistenzen der subjektiven Regelsicherheiten bei den unterschiedlichen Regelanwendern. Bei beiden Intensitätskombinationen erzielten die Durchschnittsanwender die höchsten Korrelationen (bei den gleichen: $r = .94$, $p < .001$; bei den unterschiedlichen: $r = .89$, $p < .001$), gefolgt von den Additionsanwendern (bei den gleichen: $r = .70$, $p < .001$; bei den unterschiedlichen: $r = .34$, $p = .05$). Die inkonsistenten Regelanwender waren auch bei den Einschätzungen der subjektiven Regelsicherheit innerhalb der Intensitätskombinationen am wenigsten konsistent (bei den gleichen: $r = .41$, $p < .01$; bei den unterschiedlichen: $r = .12$, ns).

Mittlere subjektive Sicherheit der prädiktiven Mischungsurteile in Abhängigkeit von der Intensitätskombination und vom Alter

Da sich im obigen Abschnitt gezeigt hatte, dass die Messwiederholungsdaten stark korrelierten, wurden die folgenden Berechnungen auf der Basis der mittleren subjektiven Sicherheitsurteile pro Intensitätskombination durchgeführt.

Bei der dreifaktoriellen Varianzanalyse nach dem Allgemeinen linearen Modell zur Überprüfung der Effekte der drei Faktoren Alter, Geschlecht und Farbinsintensitätskombination auf die mittlere Regelsicherheit der prädiktiven Regel pro Intensitätskombination zeigten sich bedeutsame Effekte des Alters ($F(4, 130) = 8.47, p < .001$), der Farbinsintensitätskombination ($F(1, 130) = 5.8, p < .05$) und eine Wechselwirkung dieser beiden Faktoren ($F(4, 130) = 2.69, p < .05$). Die Abbildung 3.2.6 stellt die Mittelwerte für jede Intensitätskombination in Abhängigkeit von der Altersgruppe dar. Die Geschlechtszugehörigkeit hatte keinen differenziellen Einfluss auf die Urteile. Eine erneute Durchführung der Varianzanalyse mit dem Faktor Reihenfolge der Versuchsdurchgänge anstelle des Geschlechtseffektors zeigte, dass die Reihenfolge weder einen direkten Haupteffekt auf die subjektiven Sicherheitsurteile hatte noch in einer Wechselwirkung mit einem anderen Faktor statistische Bedeutung erlangte.

Der Post-Hoc-Test nach Scheffé zeigte Altersunterschiede des generellen Sicherheitsurteils bei den Vergleichen der Sechs-, Acht- und Zehnjährigen mit den Erwachsenen ($p = .001, p < .001, p < .001$) sowie beim Vergleich der Zehn- und Zwölfjährigen ($p < .05$).

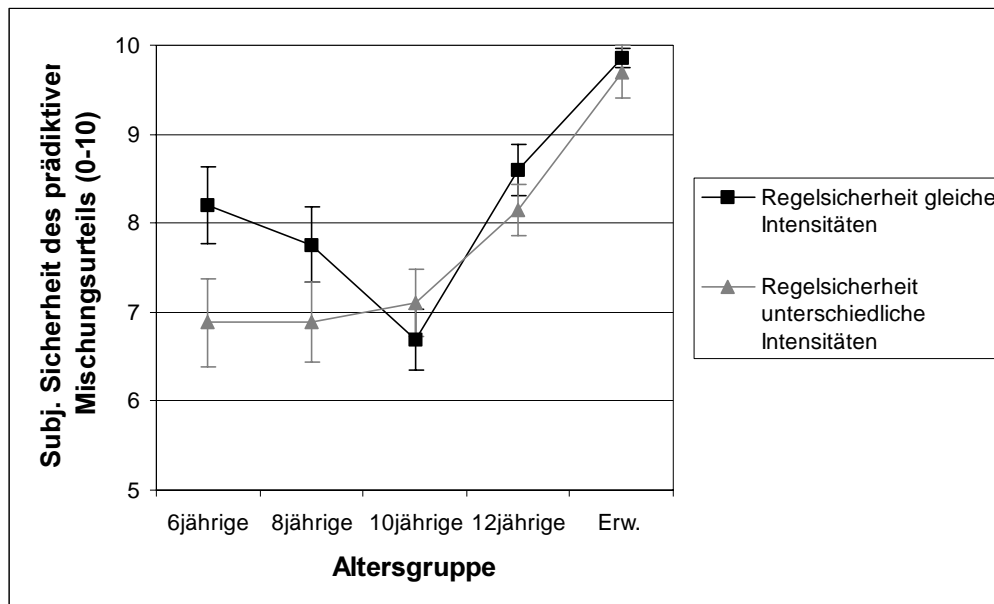


Abbildung 3.2.6. Mittlere subjektive Sicherheit und Standardfehler der subjektiven Sicherheit der prädiktiven Mischungsurteile nach Altersgruppe.

Bedingt durch das Auftreten der Wechselwirkung zwischen den Faktoren Alter und Farbintensitätskombination wurden, getrennt für die einzelnen Altersgruppen, Varianzanalysen mit dem Faktor Intensitätskombination berechnet. Der Vergleich der Mischungen gleicher und unterschiedlicher Intensitäten brachte für die Sechsjährigen einen bedeutsamen Unterschied, $F(1, 29) = 6.59, p < .05$. Wie der obigen Abbildung zu entnehmen ist, lagen die Urteile bei der Mischung gleicher Intensitäten auch bei den Achtjährigen höher als diejenigen bei der Mischung unterschiedlicher Intensitäten. Die Unterschiede erreichten die Signifikanzgrenze jedoch knapp nicht mehr, $F(1, 30) = 3.69, p = .065$. Bei den Zehnjährigen waren die subjektiven Sicherheitseinschätzungen des prädiktiven Mischungsergebnisses bei den gleichen Intensitäten leicht tiefer als diejenigen der unterschiedlichen Intensitäten. Die Zwölfjährigen und die Erwachsenen urteilten in beiden Bedingungen ähnlich, nämlich mit mehr Sicherheit bei der gleichen Intensitätskombination. Die varianzanalytischen Berechnungen bei den Zehn- und Zwölfjährigen sowie den Erwachsenen waren nicht notwendig, da schon aus Abbildung 3.2.6 zu entnehmen war, dass sich die Urteile nur gering unterschieden.

Mittlere subjektive Sicherheit der prädiktiven Mischungsurteile in Abhängigkeit von der Regel

Die univariate Varianzanalyse bei der gleichen Intensitätskombination mit den Faktoren konsistente Regelanwendung und Alter ergab höchst signifikante Effekte der konsistenten Regelanwendung ($F(3, 140) = 6.42, p < .001$) und des Alters ($F(4, 140) = 5.14, p = .001$). Der Post-Hoc-Test nach Scheffé zeigte signifikante Unterschiede in der subjektiven Regelsicherheit zwischen den inkonsistenten und den Durchschnittsregelanwendern (beobachtete mittlere Differenz = 2.38, $p < .001$) sowie zwischen den inkonsistenten und den Additionsregelanwendern (beobachtete mittlere Differenz = 1.36, $p < .05$).

Bei der unterschiedlichen Intensitätskombination ergab die Berechnung der univariaten Varianzanalyse mit den Faktoren konsistente Regelanwendung und Alter einen sehr signifikanten Effekt der konsistenten Regelanwendung ($F(3, 140) = 4.89, p < .01$). Der Faktor Alter wurde bei der Berechnung mit dem Faktor prädiktive Regelanwendung nicht mehr signifikant. Der Post-Hoc-Test nach Scheffé zeigte signifikante Unterschiede in der subjektiven Regelsicherheit zwischen den inkonsistenten und den Durchschnittsregelanwendern (beobachtete mittlere Differenz = 2.12, $p < .001$) sowie zwischen den Additions- und den Durchschnittsregelanwendern (beobachtete mittlere Differenz = 1.26, $p < .05$).

Die Mittelwerte und Standardfehler der mittleren subjektiven Sicherheitseinschätzungen der angewendeten Mischungsregel sind in Abbildung 3.2.7 in Abhängigkeit von der Art der prädiktiven Regelanwendung dargestellt.

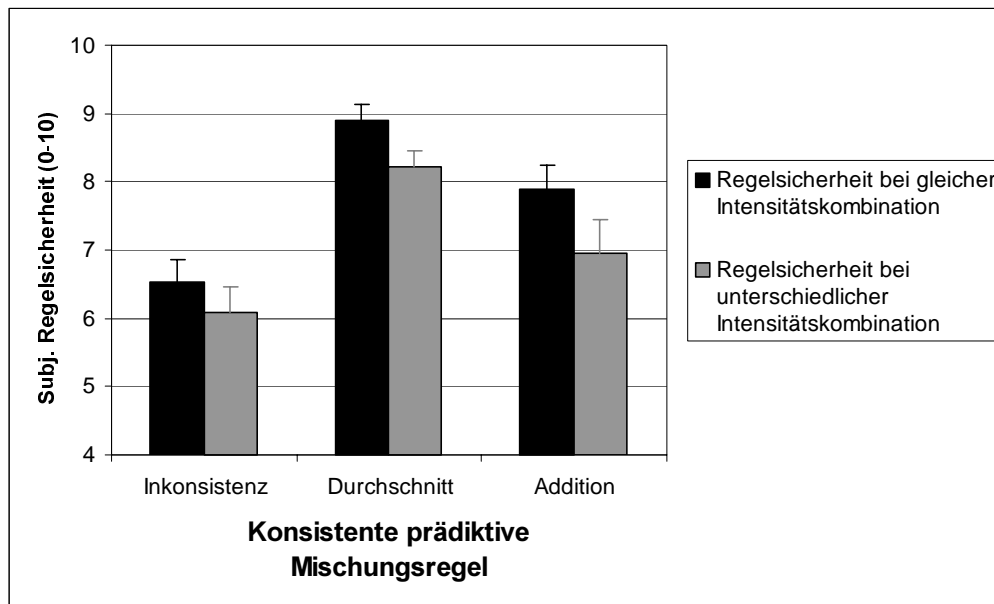


Abbildung 3.2.7. Mittelwerte und Standardfehler der subjektiven Regelsicherheit nach angewendeter Mischungsregel.

Die Betrachtung der Entwicklungsverläufe der subjektiven Regelsicherheit für jede einzelne Regel zeigte bei den Durchschnittsregelanwendern sehr ähnliche Verläufe wie bei allen Regelanwendern zusammen. Zu erwähnen bleibt, dass die Achtjährigen bei beiden Intensitätskombinationen höher waren als die Zehnjährigen. Die Addierer waren in ihrer subjektiven Sicherheit fast in jeder Altersgruppe tiefer als die Durchschnittsanwender. Die Zwölfjährigen waren sich bei beiden Intensitätskombinationen leicht sicherer als die Zehn- und Achtjährigen. Dennoch waren die zwölfjährigen Addierer deutlich unsicherer als die gleichaltrigen Durchschnittsregelanwender. Bei den inkonsistenten Regelanwendern lässt sich generell aussagen, dass diese ihre Gewissheit der Mischungsurteile stets tiefer einschätzten als die anderen Regelanwender. Ansonsten verliefen die Entwicklungsverläufe bei beiden Intensitätskombinationen unterschiedlich.

3.2.5 Diskussion

Prädiktive Mischungsregeln

Bei der Mischung gleicher Farbintensitäten zeigte der Verlauf der korrekten Durchschnittsbildung einen Anstieg mit zunehmendem Alter mit Ausnahme der Zehnjährigen, die weniger häufig korrekt antworteten als die Achtjährigen. Der Entwicklungsverlauf kann deshalb auch als annähernd u-förmig beschrieben werden. Die Häufigkeiten der

Additionsregeln hingegen blieben bei den Kindern aller Altersgruppen ziemlich konstant mit einer durchschnittlichen Häufigkeit von gut 20 Prozent Anteil pro Altersgruppe. Bei den jüngsten Kindern sowie der Altersgruppe der Zehnjährigen war der Anteil an inkonsistenten Antworten etwa so hoch wie der Anteil konsistenter Durchschnittsregeln. Die Restkategorie Subtraktionsregel wurde nur selten konsistent bei beiden Versuchsdurchgängen angewendet.

Bei der Mischung unterschiedlicher Farbintensitäten hingegen zeigte sich ein sehr deutlicher Anstieg der richtigen Durchschnittsregel mit zunehmendem Alter der Versuchspersonen. Während bei den Sechsjährigen noch vorwiegend inkonsistente Antworten gegeben wurden, betrug der Anteil der korrekten Lösungen bei den Achtjährigen schon 50 Prozent und stieg bei den Zehnjährigen auf 77 Prozent an. Komplementär dazu verliefen die Verläufe der Additionsregel und der inkonsistenten Regelanwendung. Auch bei der Mischung unterschiedlicher Farbintensitäten wurde die dritte Antwortkategorie, welche der Zentrierungsregel entsprochen hätte, kaum ausgewählt.

Der Verlauf derjenigen Versuchspersonen, welche bei beiden Intensitätskombinationen konsistent nach der Durchschnittsregel urteilten, verlief nicht durchgehend stetig ansteigend, sondern blieb bei den Acht- und Zehnjährigen auf dem gleichen Niveau von 30 Prozent, stieg bei den Zwölfjährigen deutlich auf gut 50 Prozent und bei den Erwachsenen auf 90 Prozent an. Der Entwicklungsverlauf der subjektiven Sicherheitseinschätzungen der Mischungsurteile zwischen den Acht- und Zwölfjährigen sowie den Erwachsenen war demjenigen der konsistenten Durchschnittsanwendungen in beiden Bedingungen sehr ähnlich. Diese Ergebnisse lassen vermuten, dass erst nach dem Alter von acht Jahren eine Durchschnittsregel entwickelt wird, die in verschiedenen Bedingungen angewendet wird und auch mit entsprechender Gewissheit vertreten wird. Allerdings war auch nach dem zwölften Lebensjahr noch eine Verbesserung und Verfestigung festzustellen, wie die Ergebnisse der Erwachsenen zeigten. In diesem Punkt decken sich diese neuen Resultate mit den Ergebnissen früherer Studien zur Prädiktion von Mischungsergebnissen: Sowohl bei den Forschungsarbeiten von Strauss und Mitarbeitern (v. a. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982) als auch bei den Experimenten von Jäger (1996; Jäger & Becker, 1997; Jäger & Wilkening, 2001) schätzten die meisten der Zwölfjährigen stets in verschiedenen Aufgaben das Mischungsergebnis korrekt ein.

Ein weiteres wichtiges Resultat sind die Unterschiede zwischen den beiden Intensitätskombinationen: Das Mischungsergebnis unterschiedlicher Ursprungsintensitäten wurde viel häufiger korrekt vorausgesagt als dasjenige von gleichen Ursprungsintensitäten. Dieser Unterschied war auf die Gruppe der Zehnjährigen zurückzuführen, die in der Aufgabe der Mischung von zwei gleich hellen Flüssigkeiten sogar noch schlechter abschnitten als die Sechsjährigen. Eine weitere wichtige Beobachtung war, dass die Zehnjährigen nicht häufiger nach der Additionsregel urteilten als die Acht- und Zwölfjährigen. Dies ist insofern von Bedeutung, als bei den zehnjährigen Kindern im Vergleich mit den anderen Kindern nicht eine andere Regel im Vordergrund stand, sondern ihre Leistung schlechter war.

Durch die Arbeiten von Dixon, Moore und weiteren Mitarbeitern (Ahl, Moore & Dixon, 1992; Dixon & Moore, 1996; Moore, Dixon & Haines, 1992) wurden verschiedene Teilstrategien, die zur korrekten Durchschnittsbildung vorhanden sein müssen, identifiziert. Bei den Aufgaben beider Intensitätskombinationen käme die Teilstrategie des sogenannten *Range Principle* – die resultierende Temperatur der Mischung zweier Komponenten muss innerhalb des Bereichs der beiden ursprünglichen Temperaturen liegen – zum Tragen. Die Antworten bei der gleichen Intensitätskombinationen zeigten, dass diese Teilstrategie auch bei Kindern im Alter von zehn Jahren deutlich verletzt wurde. Bei den Aufgaben der beiden Intensitätskombinationen kamen aber auch unterschiedliche Teilstrategien zum Einsatz: Bei der Mischung gleicher Intensitätskombination spielte das *Equal-Intensities-Equal principle*, das sich auf die Mischung von Komponenten mit gleich hohen intensiven Grössen bezieht, eine Rolle. Nach diesem Prinzip findet beim Beispiel von Temperaturmischungen kein Temperatúraustausch statt, die Temperatur bleibt konstant. Bei der Mischung von Flüssigkeiten mit unterschiedlichen intensiven Grössen hingegen hat das Prinzip *Above-Below* eine wichtige Funktion. Hier richtet sich der Fokus nur auf die Veränderung der ursprünglichen intensiven Grösse einer Komponente, welche sich immer in Richtung der zugegebenen intensiven Grösse verändert. Wird also zu einer dunkleren Flüssigkeit eine hellere dazugegeben, so wird die resultierende Flüssigkeit heller werden als die dunklere der beiden. Obwohl das unterschiedliche Lösen bei den beiden Intensitätskombinationsbedingungen die Berücksichtigung verschiedener Teilstrategien im Sinne von Dixon und Moore beinhaltet, können die unterschiedlichen Befunde in der vorliegenden Studie mit diesem Ansatz nicht erklärt werden. Es kann nur bestätigt werden, dass auch scheinbar einfachere Prinzipien

wie das *Range Principle* und das *Equal-Intensities-Equal principle* noch von Kindern im Alter von zehn Jahren verletzt wurden.

Ergebnisse früherer Untersuchungen mit Mischungsaufgaben bei intensiven Grössen deuteten bereits darauf hin, dass die kindlichen Lösungsversuche stark von den Aufgabensituationen abhingen. Diese Vermutung bestätigte sich auch durch die vorliegende Untersuchung durch die folgenden Ergebnisse: Erstens wurden die Aufgaben der beiden Intensitätskombinationen unterschiedlich gut gelöst. Zweitens führte die nur geringe Variation der Aufgabensituation im Vergleich zu früheren Untersuchungen zu gegensätzlichen Resultaten. Die unterschiedlichen Resultate werden zunächst als Ausdruck des adaptiven Denkens von Kindern verstanden und interpretiert.

Betrachtung der Ergebnisse im Lichte des Ansatzes des adaptiven Denkens

Der Ansatz des adaptiven Denkens stellt die Anpassungsversuche der kindlichen Denkvorgänge an wechselnde situative Einflüsse in den Vordergrund. Wenn Kinder noch nicht über das notwendig physikalische Wissen verfügen, um die Aufgabe zu lösen, dann versuchen sie mithilfe ihres Vorwissens und Hinweisreizen aus der Aufgabendarbietung eine Lösungsregel zu finden.

So drängt sich beim vorliegenden Experiment zunächst die Frage auf, weshalb die Kinder die Mischungsergebnisse von zwei Flüssigkeiten, die unterschiedlich hell waren, so gut vorhersagen konnten. Ab dem Alter von 10 Jahren waren nicht nur fast 80 Prozent der Mischungsurteile der Kinder konsistent korrekt, sondern auch die Additionen der Ursprungsintensitäten waren auf dem tiefen Niveau von ca. 10 Prozent der Antworten.

Da der Grund für das gute Abschneiden bei der unterschiedlichen Intensitätskombination möglicherweise in einem kleinen Detail der Aufgabendarbietung steckt, lohnt sich eine genaue Analyse der Aufgabenstellungen. Hilfe bieten hierzu auch frühere Untersuchungen mit sehr ähnlichen Aufgabenstellungen. Auch bei Experimenten von Jäger (Jäger, 1996; Jäger & Becker, 1997; Jäger & Wilkening, 2001) mussten Mischungsergebnisse von zwei Flüssigkeiten mit unterschiedlichen Farbintensitäten vorausgesagt werden. Bei allen vorgängigen Experimenten waren aber die Aufgabenstellungen für die Kinder offensichtlicher weniger geeignet, die korrekte Lösung zu finden. Allen früheren Experimenten gemeinsam war, dass die Zehn- und z. T. die zwölfjährigen Kinder deutlich

weniger gut waren als im vorliegenden Experiment. Es könnte also sein, dass eine kleine Variation der früheren Experimente daran beteiligt war, den Kindern zu einer besseren Lösungsstrategie zu verhelfen.

Sowohl bei den früheren Experimenten von Jäger als auch beim vorliegenden Experiment wurde die intensive Grösse in Form von Farbhelligkeiten direkt wahrnehmbar dargeboten. Unterschiede bestanden aber im Mischungsvorgang und in der Erfassung der Mischungsurteile. Antworten, die mit der Wahlaufgabe erfasst werden, gelten im Vergleich zu Antworten, die auf einer fein abgestuften Skala gemacht werden müssen, als die weniger präzisen. Da aber beim vorliegenden Experiment, bei dem die besseren Einschätzungen erfolgten, die Wahlaufgabe verwendet wurde, ist es schwer vorstellbar, wie die Art der abhängigen Variablen entscheidend für die Mischungsurteile hatte sein können.

Der andere – vermutlich entscheidendere – Unterschied betraf die Vorstellung des Zusammenleerens der beiden Flüssigkeiten. Während bei den früheren Studien die beiden Flüssigkeiten gleichzeitig in ein drittes Glas geleert werden sollten, war im vorliegenden Experiment kein drittes Glas vorhanden, da die eine Flüssigkeit zu der anderen Flüssigkeit dazu geleert werden sollte. Vielleicht lag das entscheidende Detail darin, dass immer die hellere Flüssigkeit in die dunklere geleert werden sollte. Gleichzeitig wurde das Kind in der Instruktion noch auf eine Analogie im Alltag, nämlich dem Mischungsvorgang von Sirup, aufmerksam gemacht. Da sich in Vorversuchen gezeigt hatte, dass die Kinder mit dieser Analogie die Aufgabe gut lösen konnten, wurde sie auch für die Versuchsdurchführung der Hauptstudie verwendet. Gleichzeitig kann man annehmen, dass der Mischungsvorgang, bei dem eine helle Flüssigkeit – die an leicht gefärbtes Wasser erinnerte – in eine dunkle Flüssigkeit den Kindern aus dem Alltag vertraut war. Es ist auch denkbar, dass ein kognitives Schema im Sinne des mentalen Balancierens (Anderson, 1981, 1987) aktiviert werden konnte. Bei diesem Mischungsvorgang konnten sich die Kinder vielleicht besonders gut vorstellen, wie sich die helle Flüssigkeit in der dunkleren verteilt und diese dadurch etwas heller wird. Währenddessen könnte ein gleichzeitiges Zusammenleeren der Flüssigkeiten aufgrund der weniger direkten Analogie dieses kognitive Schema nicht in gleichem Ausmass aktiviert haben. Denkbar ist auch, dass dieser Vorgang bei der Aufgabendarbietung von Jäger (1996; Jäger & Wilkening, 2001) höhere kognitive Anforderungen an das Kind gestellt hatte,

womit zumindest die sehr selten korrekten Antworten der Sechsjährigen beim gleichzeitigen Zusammenleeren in ein drittes Glas erklärt werden könnten.

Es ist weiter vorstellbar, dass dieses kognitive Schema des mentalen Ausbalancierens nur bei zwei unterschiedlichen Farbintensitäten zum Tragen kommen konnte. So könnten die schlechteren Mischungsurteile bei der gleichen Intensitätskombination im vorliegenden Experiment erklärt werden. Möglicherweise wird die Vorstellung vom Ineinanderfließen bei zwei gleich hellen Flüssigkeiten automatisch gestoppt, da man sich den Austausch nicht visuell vorstellen kann. Im Falle eines fehlenden vertrauten Schemas könnte das *direct functional reasoning* im Sinne von Strauss (1982a) oder die *general-purpose-adding-rule* von Anderson und Cuneo (1978a) wieder ins Spiel kommen.

Bei Jägers untersuchten Voraussagefähigkeiten bei der Mischung von zwei gleich intensiven Flüssigkeiten (1996, Experiment 1) kehrte sich das Bild wieder: Die jüngeren Kinder urteilten viel eher nach der Durchschnittsregel und dies bereits die Sechsjährigen zu etwa 75 Prozent. Eigentlich hätten Jägers unterschiedliche Verläufe bei den beiden Intensitätskombinationen schon dafür gesprochen, dass unterschiedliches Funktionswissen aktiviert wurde, was aber von Jäger nicht diskutiert wurde. Zur Klärung der verschiedenen Resultate bei den Mischungsvorgängen gleicher Intensitäten stellt sich erneut die Frage, ob durch die geringe Aufgabenvariation verschiedene Schemata aktiviert worden waren. Möglich wäre, dass beim gleichzeitigen Zusammenleeren bei Jäger, anstelle des kognitiven Modells des Ausbalancierens ein anderes Modell aktiviert wurde, das auch bei den jüngsten Kindern zu guten Voraussagen führte. Denkbar ist etwa die Vorstellung der Identitätserhaltung: Werden zwei gleich helle Flüssigkeiten gleichzeitig in ein drittes Glas geleert, so bleiben sie auch gleich. Diese Identitätserhaltung benötigt kein Antizipieren des konkreten Mischungsvorganges und deshalb auch keine höhere Verarbeitungskapazität, weshalb auch die Sechsjährigen diese Aufgabe gut lösen konnten. Schon Stavy, Strauss und andere Mitarbeiter (Stavy et al., 1982) hatte bei den erfassten Erklärungen bei Kindern in diesem Alter festgestellt, dass sie durch Identitätserhaltung zur korrekten Lösung kamen.

Noch komplizierter wird die Sachlage, wenn man die Ergebnisse der Untersuchungen der Gruppe um Strauss (v. a. Stavy, 1981; Stavy et al., 1982; Strauss & Stavy, 1982), die ebenfalls die Prädiktion des Mischungsergebnisses von zwei gleichen Intensitäten be-

inhalten, hinzunimmt. Bei diesen Studien war das vorherrschende Ergebnis das häufige Addieren der intensiven Grösse. Dies, obwohl auch in diesen Aufgaben – wie bei Jäger – die beiden Flüssigkeiten gleichzeitig in ein drittes leeres Glas geleert werden sollten.

Jäger hatte sich die verbesserten Lösungsstrategien bei der gleichen Intensitätskombination – ohne die schlechtere Leistung bei der unterschiedlichen Intensitätskombination dabei zu berücksichtigen – mit der direkt visuell wahrnehmbaren Darbietung der intensiven Grösse erklärt. Zur Erinnerung: Strauss hatte die intensive Grösse mit Anzahl Kerzen (Temperatur) und Anzahl Zuckerstücke oder mit Zucker gefüllte Löffel (Süsse) dargeboten. Es wird an dieser Stelle nach wie vor für möglich gehalten, dass dieser Hinweisreiz in Form von quantitativen Angaben so stark war, dass die intensive Grösse wie eine extensive behandelt wurde. War das gute Abschneiden vorher mit der Identitätserhaltung erklärt worden, so scheint diese bei der Darbietung der intensiven Grössen mit quantitativen Angaben fast nicht möglich.

Jäger hatte den Unterschied zudem noch an der Paarvergleichsmethode bzw. Wahlaufgabe festgemacht. Die vorgegebenen Antwortkategorien würden eine direkte funktionale Beziehung (*direct functional reasoning*) zwischen Farbintensität und der Vergleichsflüssigkeit, die halbes Volumen hatte, im Sinne von „mehr Flüssigkeit“ = „mehr Farbe“ bzw. „dunkler“ in den Vordergrund stellen. Es ist denkbar, dass die Wahlaufgabe auch im vorliegenden Experiment das Addieren der Farbintensitäten verstärkt hatte. Dies wird v. a. dann vermutet, wenn kein kognitives Schema zur Verfügung steht und deshalb auf andere Hinweisreize zurück gegriffen werden muss.

Zusammengefasst lässt sich sagen, dass das Urteilsverhalten der Kinder erst ab zwölf Jahren relativ unbeeinflusst von der Aufgabenstellung ausfiel. Dennoch urteilten noch nicht alle Zwölfjährigen korrekt, ein fast perfektes Verständnis der kognitiven Durchschnittsbildung zeigte sich erst bei den erwachsenen Studierenden. Bei den Kindern bis zehn Jahren schien das Denken adaptiv zu sein, da sich bei den jüngeren Kindern eine hochgradige Abhängigkeit von der Aufgabedarbietung anzeigte.

Rückblickend betrachtet bestand die Aufgabe der Interpretation der Ergebnisse – insbesondere der jüngeren Kinder – in einer detaillierten Analyse der spezifischen Aufgabenstellungen von verschiedenen Untersuchungen. Bereits geringe Aufgabenvariationen

konnten potenziell als Auslöser für die Aktivierung verschiedener kognitiver Schemata infrage kommen.

Die Resultate wurden dahingehend interpretiert, dass im Falle einer Aktivierung eines kognitiven Schemas bzw. Funktionswissens auch die jüngeren Kinder viel weniger Schwierigkeiten hatten, die Aufgabe korrekt zu lösen. Wenn ein starkes kognitives Schema zur Verfügung stand, fielen die Antworten relativ eindeutig aus: In der vorliegenden Untersuchung wird bei der unterschiedlichen Intensitätskombination die Aktivierung eines Schemas im Sinne des mentalen Balancierens vermutet. Beim Experiment von Jäger (1996) mit der Kombination gleicher Ursprungsintensitäten war die Vorstellung einer Identitätserhaltung, die auch Sechsjährige zur richtigen Lösung führten, eine plausible Denkstrategie. Falls das Schema zur korrekten Lösung führte, wurde dieses mit zunehmendem Alter auch stetig häufiger aktiviert.

Konnte aber aufgrund der erschwerten Aufgabendarbietung kein Schema aktiviert werden, fielen die Urteile uneinheitlicher aus. Es ist nahe liegend, dass die Aufgabe auf andere Hinweisreize abgesucht wird, die auf eine Lösung hinführen könnten. Die ersichtlichsten Hinweisreize waren die quantitative Darbietung der intensiven Grösse und die Wahlaufgabe. Beide führten vorwiegend zu einer Addition der intensiven Grösse. Bei den jüngeren Kindern wurde auch vermutet, dass Aufgabenstellungen, die hohe kognitive Anforderungen an die Versuchspersonen stellten, zu vielen inkonsistenten Urteilen führten.

Aufgrund der unterschiedlichen Altersverläufe der korrekten Mischungsurteile und des stark variierenden Ausmasses von additiven und inkonsistenten Urteilen bei den Kindern bis zum Alter von zehn Jahren war kein einheitlicher Entwicklungstrend von Mischungsregeln zu beobachten. Die Schlussfolgerung von Jäger (1996), dass sich die kognitive Durchschnittsbildung in Schritten von inkonsistenten Urteilen über additive Mischungsregel zur korrekten Durchschnittsbildung hin entwickle, scheint deshalb nicht mehr haltbar.

Weiter zeigte sich sowohl bei Jägers Experiment (1996; Experiment 1, bei beiden Intensitätskombinationen) als auch bei der vorliegenden Untersuchung (unterschiedliche Intensitätskombination) eine Tendenz zu einem u-förmigen Entwicklungsverlauf. Diese

Tendenz war jeweils auf die Altersgruppe der Zehnjährigen zurückzuführen. In einem Fall addierten sie häufiger, im anderen Fall urteilten sie häufiger inkonsistent als die Achtjährigen. Die Verteilungsunterschiede der konsistent korrekten Regeln und der inkonsistenten Regeln bei den Sechs- bis Zehnjährigen bei der vorliegenden Untersuchung waren allerdings nicht vom Alter abhängig. Der Rückgang an korrekten Mischungsurteilen war nur bei der unterschiedlichen Intensitätskombination im Experiment von Jäger (1996) sehr deutlich. Ein allfälliger Anspruch auf Universalität des u-förmigen Entwicklungsverlaufs der kognitiven Durchschnittsbildung (Strauss, 1982c) müsste aufgrund der vorliegenden Resultate verworfen werden. Dennoch weisen auch die Daten von Phase 1 auf erstaunlich ausgeprägte Schwierigkeiten bei der Mischungsvoraussage bei den Zehnjährigen zu bestehen: Erstens ihre waren Urteile in der Bedingung der gleichen Intensitätskombinationen erstaunlich häufig inkonsistent. Fasst man das adaptive Denken als eine intuitive Lösungsstrategie auf, so könnte eine Interaktion von intuitivem und analytischem Denken (vgl. Abschnitt zur Durchschnittsbildung auf der Basis intuitiven vs. analytischen Denkens, S. 23f.) bei den Zehnjährigen vorgelegen haben. Die besseren Einschätzungen der Achtjährigen bei der gleichen Intensitätskombination könnten durch ein ungehemmteres intuitives bzw. adaptives Problemlösen sprechen. Allerdings scheint diese Blockierung nur einzutreten, wenn kein starkes kognitives Schema vorhanden ist (vgl. die hohen korrekten Urteile bei der gleichen Intensitätskombination). Die Zwölfjährigen hingegen könnten eine zunehmend analytische Denkweise zu den relativ aufgabenunabhängigen Urteilen und höheren Gewissheiten ihrer Urteile geführt haben. Ein zweiter Hinweis auf Urteilsschwierigkeiten der Zehnjährigen ist den Einschätzungen der subjektiven Sicherheiten der Mischungsurteile zu entnehmen. Allerdings entspricht ein Sicherheitseinbruch bei den Zehnjährigen auch einer allgemeinen Tendenz, die bereits in anderen Studien gefunden wurde und weniger mit der Art oder Verfügbarkeit einer Denkstrategie bei den jüngeren Kindern denn als einer generellen Unfähigkeit zur Äusserung von Unsicherheiten erklärt wird (vgl. Allgemeine Diskussion, S. 231ff.).

Die Ergebnisse aus Phase 1 verdeutlichten, wie stark die Lösungen in Aufgabenstellungen mit nur wenig Unterschied variieren können. Deren Interpretation führte zur Verwerfung früherer Hypothesen, die aufgrund einer einzigen Aufgabendarbietung aufgestellt wurden. Aus dieser Sicht wird deutlich, wie beschränkt die Ergebnisse von Unter-

suchungen mit nur einer Methode sein können und allgemein gültige Schlussfolgerungen nicht haltbar sind.

Subjektive Sicherheit der Mischungsurteile vor dem Mischungsvorgang

Die subjektiven Sicherheitseinschätzungen des unmittelbar vorher geäußerten Mischungsurteils zeigten bei den jüngeren Kindern grosse interindividuelle Streuungen innerhalb der Altersgruppen. Diese Streuungen wurden aber mit zunehmendem Alter geringer. Gleichzeitig nahm aber die Gewissheit der Regel mit steigendem Alter zu. Der Entwicklungsverlauf zeigte erst einen bedeutsamen Anstieg an Regelsicherheit bei den zwölfjährigen Kindern. Die Gewissheit ihrer Regel war bei den Zwölfjährigen zwar schon im oberen Bereich der Skala, bei den Erwachsenen stieg sie aber nochmals bis auf das Maximum an.

Die Abhängigkeit der subjektiven Sicherheit von der Art der angewendeten Mischungsregel erwies sich als hypothesenkonform: Bei beiden Intensitätsbedingungen hatte die Art der Mischungsregel einen bedeutsamen Einfluss auf die Regelsicherheit in der folgenden Reihenfolge mit abnehmender Gewissheit: Durchschnitt > Addition > Inkonsistenz. Die Tatsache, dass die Additionsregelanwender sich in ihrem Mischungsurteil sicherer waren als die inkonsistenten Regelanwender verstärkt die Vermutung, dass auch sie – nebst den Durchschnittsregelanwendern – bei der Lösungsfindung eine Strategie verfolgten.

Weiter zeigte sich eine Abhängigkeit bei der Gewissheit der Regel von der Intensitätskombination. Die Farbintensitätskombination der zu mischenden Flüssigkeiten hatte insgesamt und in Abhängigkeit von der Altersgruppe einen bedeutsamen Effekt auf die Regelsicherheit. Besonders bemerkbar machte sich der Unterschied bei den jüngeren Kindern. Abbildung 3.2.7 zeigte deutlich, dass die subjektive Sicherheit bei der gleichen Intensitätskombination bei den verschiedenen Regelanwendern durchgängig höher war als bei der unterschiedlichen Intensitätskombination.

Insgesamt und v. a. bei den jüngeren Kindern wurde beobachtet, dass die Mischungsurteile der unterschiedlichen Intensitätskombination unabhängig von der Regelanwendung weniger sicher vertreten werden konnten als diejenigen der gleichen Intensitätskombination. Dies kann als Hinweis dahingehend verstanden werden, dass den Mischungsur-

teilen – und dies v. a. bei den jüngeren Kindern – möglicherweise andere kognitive Prozesse zu Grunde lagen. Die postulierte ausgeprägte Aktivierung des kognitiven Schemas des mentalen Ausbalancierens bei der unterschiedlichen Intensitätskombination führte aber nicht zu grösserer Gewissheit als die Denkstrategie bei der gleichen Intensitätskombination.

3.3 Magic experiment

3.3.1 Fragestellung

Die Befunde von Studien mit der Methode des *extinction paradigm* bzw. *magic experiment* hatten gezeigt, dass mit dieser Methode Aussagen über die Stabilität, mit der eine Person ein Konzept oder eine Regel „hat“, gemacht werden können. Es wird erwartet, dass erst ein vollständig ausgereiftes Verständnis einer Regel maximale Stabilität in verschiedenen Aufgabensituationen ermöglicht. Abweichungen von einer hohen Stabilität werden als noch nicht ausgereiftes Verständnis verstanden. Es wird erwartet, dass die Stabilität der Durchschnittsregel mit zunehmendem Alter ansteigt und bei den Erwachsenen maximal ist. Nebst der Stabilitätsentwicklung der korrekten Regel interessiert aber auch die Stabilität der fehlerhaften Intensivierungsvorstellung. Ist diese genauso oder weniger stabil wie die korrekte Regel? Nimmt die Stabilität mit zunehmendem Alter ab? Mit der Verwendung des *magic experiment* wird ein Beitrag zur Stabilitätsentwicklung der korrekten Durchschnittsregel und derjenigen der falschen Additionsregel erhofft.

Nebst der Stabilitätshypothese aus der Tradition des *extinction paradigm* bzw. *magic experiment* können durch die Anwendung des *magic experiment* auch im Hinblick auf den Ansatz des adaptiven Denkens Hypothesen generiert werden. Gemäss den Studien mit dem *violation of expectation-paradigm* (z. B. Baillargeon, 1986, 1993, 1995) sowie der Studie von Lindsay und Creedon (1985) erlauben die kindlichen Reaktionen auf physikalisch mögliche und unmögliche Ereignisse einen Rückschluss auf das vorhandene Verständnis des manipulierten physikalischen Prinzips. Es ist anzunehmen, dass die Beurteilung des Mischungsverhaltens von intensiven Grössen nach Beobachtung physikalisch möglicher und unmöglicher Mischungsvorgänge für die Kinder eine leichtere Aufgabensituation im Vergleich zur prädiktiven Aufgabe darstellt.

Ausserdem wird durch Variation der Intensitätskombinationen der Frage nachgegangen, ob die Kinder bei verschiedenen Intensitätskombinationen der beiden Ursprungsflüssigkeiten gleichermassen vorgehen oder ob sie diese verschieden behandeln. Dieser Aspekt wurde in der bisherigen Forschung wenig beachtet und diskutiert. Aufgrund der Vorarbeit von Jäger (1996, Experiment 1) könnten sich Unterschiede ergeben, Erklärungsver-

suche hierzu wurden aber keine angestellt. Dieser Aspekt hat also in dieser Untersuchung primär explorativen Charakter.

Bei der Operationalisierung der kindlichen Reaktionen auf die verschiedenartigen Mischungsergebnisse hatten die Ergebnisse der Studie von Chandler und Lalonde (1994) entscheidenden Einfluss. Da bei der Präsentation gezauberter Ereignisse durch eine Versuchsleiterin in der Rolle einer Zauberin die geringsten Versuchsleiter-Effekte zu erwarten sind, wurde das Experiment in einen Zauberkontext gestellt. Um zu erfassen, für wie realistisch die Versuchspersonen die jeweiligen Mischungsergebnisse hielten, wurden sie gefragt, wie fest bei den Mischungen gezaubert bzw. getrickst worden war (Zauberskala von 0 bis 10). Wicht die Mischung von der Erwartung ab, sollte sie als gezaubert eingeschätzt werden. Entsprach die Mischung aber exakt der Erwartung, so war sie nicht gezaubert. Diese Zauberskala sollte eine Einschätzung ohne die Notwendigkeit eines bewussten Umgangs ermöglichen. Mit der elfstufigen abhängigen Variablen waren zudem differenzierte Urteile möglich. Durch die Erfassung der Reaktionen auf beide Mischungsergebnisse konnte die Unterscheidungsfähigkeit mit der Zauberskala (Differenz zwischen der Einschätzung physikalisch möglichem und unmöglichem Mischungsergebnis) gemessen werden. Je diskrepanter die beiden Einschätzungen ausfielen (beachte Vorzeichen!), umso eher konnte eine Mischungsregel unterstellt werden. Somit konnte also auch erfasst werden, welche Mischungsregel im Setting des *magic experiment* vermutet werden konnte. Damit war durch einen direkten Vergleich mit der prädiktiven Mischungsregel auch eine Aussage über die Stabilität dieser Regel möglich.

Es liegt bisher noch keine Studie vor, die zur Realitätseinschätzung von physikalisch möglichem und unmöglichem Ereignissen eine abgestufte Zauberskala verwendet hatte. Deshalb sollte auch der Frage nach der Validität dieser Rating-Skala nachgegangen werden. Verschiedene Untersuchungen (z. B. Chandler & Lalonde, 1994; Rosengren & Hickling, 1994) hatten gezeigt, dass Kinder ab dem Alter von drei Jahren einfache manipulierte Ereignisse bereits ausschliesslich den Kategorien „Zaubertricks“ oder „Magie“ zuordneten und dabei gegenüber der Kategorie „physikalisch möglich“ abgrenzten. Allerdings gehörten bei einigen Fünfjährigen die so genannten aussergewöhnlichen Transformationen, die auf chemischer Basis zu Stande kamen, bereits zur Kategorie „physikalisch“ bzw. – korrekter ausgedrückt – „chemisch möglich“. Die aussergewöhnlichen Transformationen von Rosengren und Hickling (1994) gehörten demnach – viel-

leicht weil sie mit diesen Transformationen mehr vertraut waren als mit den Zaubertricks – nicht mehr zu den klassischen Zaubertricks. Es wurde aber trotzdem nicht erwartet, dass die Farbmischungstricks bei einer Einschätzung zwischen der Dichotomie gezaubert/getrickst vs. nicht-gezaubert zu den nicht-gezauberten Ereignissen gezählt würden. Die Dichotomie Zaubern – Nicht-Zaubern sollte eigentlich klar suggerieren, dass zwischen nur scheinbar physikalisch Möglichem und wirklich physikalisch Möglichem unterschieden werden sollte. Auf die Frage nach der inhaltlichen Validität wird am Ende des Auswertungsteils (S. 145ff.) eingegangen.

Um den bewussten Umgang mit der prädiktiven Regel zu überprüfen, wurde gefragt, wie sicher sie nach der Beobachtung der einzelnen Mischungsresultate bezüglich ihrer vorher genannten Mischungsvoraussage waren. Hier mussten sie auf einer elfstufigen Skala von –5 für die sichere Verwerfung ihrer vorherigen Voraussage bis +5 für die sichere Bestätigung ihrer prädiktiven Voraussage des Mischungsresultates einschätzen. War die Versuchsperson nach Beobachtung des Mischungsresultates aber verunsichert und wusste nicht mehr, welches Resultat im Normalfall zu erwarten war, hatte sie auf der Skala den Wert null anzugeben. Bei dieser Einschätzung war bereits vor Beginn der Versuchsdurchführung fraglich, ob auch die jüngeren Kinder zu einer Einschätzung in der Lage waren, da der bewusste Umgang mit der Regel metakognitive Fähigkeiten zur Einsicht ins eigene Denken voraussetzt. Generell wurde vermutet, dass die vollständige Gewissheit einer Regel erst am Ende des Stabilisierungsprozesses erlangt wird. Es wurde also im Vergleich mit der Zauberskala erhobenen Stabilität eine Entwicklungsverzögerung erwartet. Mit der Erfassung des bewussten Umgangs mit der prädiktiven Regel nach Beobachtung widersprüchlicher Resultate wurde auch angenommen, dass sich ein allfälliger kognitiver Konflikt in einer tieferen Regelsicherheit äussern würde. Hierzu stand die Frage im Zentrum, ob die Anwender der falschen Additionsregel einem größeren Konflikt ausgesetzt waren als die korrekten Durchschnittsregelanwender. Da mit dieser Rating-Skala auch eine Verwerfung ihrer ursprünglichen Regel erfasst werden konnte, war es auch möglich, eine Aussage über einen allfälligen *cognitive change* zu machen.

Spezifische Fragestellungen zum *magic experiment*

Einschätzung der Mischungsergebnisse auf der Wahrnehmungsebene

Unmittelbar nach der Beobachtung des Mischungsvorganges musste zuerst beurteilt werden, was aus der ursprünglichen Ausgangsintensität des einen Glases, in das die Flüssigkeit des anderen Glases geleert wurde, geworden war. Dabei waren jeweils drei Einschätzungen möglich: Es war gleich geblieben, es war dunkler geworden oder es war heller geworden. Zur Vereinfachung der Einschätzung wurde jeweils ein Referenzglas mit einer Farbintensität, die der ursprünglichen Ausgangsintensität entsprach, neben die Mischungsflüssigkeit gestellt. Aufgrund dieser Beurteilungen konnten die Fehler der Einschätzungen der Mischungsergebnisse auf der Wahrnehmungsebene bei den einzelnen Versuchsdurchgängen bestimmt werden. Aus methodischer Sicht war es wichtig, sicherzustellen, dass die korrekten Einschätzungen der Mischungsergebnisse auf der Wahrnehmungsebene nicht durch Raten entstanden sein konnten. Von eher peripherer Bedeutung war die Frage, ob sich Altersunterschiede in der Häufigkeit der Abweichungen zeigten und ob allenfalls die prädiktive Mischungsregel einen Einfluss auf die Wahrnehmungseinschätzung hatte.

Realitätseinschätzungen anhand der Zauberskala

Mischungsart: Hatte die Art des Mischungsergebnisses Einfluss auf die Einschätzungen mit der Zauberskala? Es wurde erwartet, dass die physikalisch korrekten Mischungsergebnisse als weniger gezaubert bzw. als realistischer eingeschätzt wurden als die physikalisch unkorrekten. Zudem wurde vermutet, dass der Einfluss der Mischungsart mit zunehmendem Alter deutlicher wurde.

Intensitätskombination: Zeigte sich ein Unterschied zwischen den beiden Intensitätskombinationen bei der Einschätzung mit der Zauberskala? Da sich in Experiment 1 von Jäger (1996) und in der vorliegenden Untersuchung in Phase 1 zeigte, dass die Mischungsergebnisse der Kombinationen gleicher und unterschiedlicher Ursprungsintensitäten der Prädiktionsbedingung verschieden gut vorausgesagt werden konnten, besteht die Vermutung, dass auch beim *magic experiment* Unterschiede zwischen den beiden Intensitätskombinationen zum Ausdruck kamen.

Prädiktive Regelanwendung: Ein zentrales Ziel der Phase 2 bestand in der Überprüfung, ob sich die verschiedenen Regelanwender bei der Einschätzung physikalisch korrekter und unkorrekter Mischungsresultate unterschieden. Theoretisch wurden sowohl den Anwendern der Additionsregel als auch der Durchschnittsbildung je der Erwartung entsprechende als auch widersprechende Mischungsresultate präsentiert. Eine zentrale Frage lautete: Behandeln die verschiedenen konsistenten bzw. inkonsistenten Regelanwender die Mischungsresultate gleich oder verschieden? Werden im *magic experiment* die Mischungsresultate bei den verschiedenen Regelanwendern gleich stabil entsprechend ihrer prädiktiven Regel eingeschätzt? Konkret wurden folgende Unterschiede in den Einschätzungen der beiden Mischungsarten zwischen den verschiedenen Regelanwendern aus Phase 1 erwartet: Die korrekten konsistenten Regelanwender sollten die physikalisch korrekten Mischungsresultate als weniger gezaubert einschätzen als die physikalisch unkorrekten und die prädiktiven Additionsregelanwender sollten die physikalisch unkorrekten Mischungsresultate als weniger gezaubert als die physikalisch korrekten einschätzen. Ferner interessierte auch, ob sich Entwicklungsunterschiede bei verschiedenen „guten“ Anwendern bei den Einschätzungen der beiden Mischungsresultate mit der Zauberskala zeigten und ob diese bei beiden Mischungsarten gleich verliefen.

Da beim Probeversuchdurchgang die Reaktion auf die physikalisch unkorrekte Mischung von zwei farblosen Flüssigkeiten mit der Zauberskala erhoben wurde, konnte diese Reaktion mit denjenigen auf die physikalisch unmöglichen und möglichen Mischungen der vier Versuchsdurchgänge verglichen werden. Hier interessierte insbesondere die Frage, ob das physikalisch unkorrekte Mischungsresultat bei der Mischung von zwei farblosen Flüssigkeiten – generell und in Abhängigkeit vom Alter – als mehr gezaubert eingeschätzt wurde als die physikalisch unkorrekten Mischungsresultate von farbigen Flüssigkeiten.

Diskrimination zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten

Mischungsart: Wurden die beiden Mischungsarten mit zunehmendem Alter besser unterschieden? Da die korrekte Mischungsregel mit dem Alter erworben wird, wurde erwartet, dass auch die Diskriminationsfähigkeit zwischen physikalisch korrekten und unkorrekten mit zunehmendem Alter besser wurde.

Intensitätskombination: Entsprechend der Resultate in der prädiktiven Bedingung wurde erwartet, dass sich die Diskriminierungsfähigkeit bei der gleichen und unterschiedlichen Intensitätskombination unterschieden hatte. Allerdings war schwierig vorauszusagen, bei welcher Intensitätskombination besser diskriminiert werden sollte, da die Ergebnisse aus Jägers Studie (Experiment 1, 1996) und die Resultate dieser Untersuchung von Phase 1 widersprüchlich waren. Es wurde aber vermutet, dass sich dieser Unterschied mit zunehmendem Alter aufhob.

Prädiktive Regelanwendung: Eine der wichtigsten Hypothesen dieses Experimentes bestand in der Annahme, dass sich die verschiedenen Regelanwender in ihrer Diskrimination der beiden Mischungsresultate grundlegend unterschieden. Es wurde erwartet, dass die Durchschnittsbilder der Phase 1 im positiven Sinne diskriminierten, d. h. dass sie die physikalisch unkorrekten Mischungsresultate als mehr gezaubert eingeschätzt hatten als die korrekten Mischungen und die Additionsregelanwender im negativen Sinne, d. h. dass die physikalisch unkorrekten Mischungsresultate weniger gezaubert waren als die korrekten Mischungen. Bei den Durchschnittsregelanwendern wurde erwartet, dass die positive Unterscheidung mit zunehmendem Alter ausgeprägter wurde. Bei den Additionsregelanwendern war es schwieriger eine Prognose abzugeben: Handelten sie mit zunehmendem Alter mehr ihrer prädiktiven Regel entsprechend oder wurde diese vermehrt aufgegeben?

Mit dem Mass Diskriminierungsfähigkeit war es möglich, die Einschätzungen in korrekte (positive Werte), unkorrekte (negative Werte) und indifferente (Differenzwert von 0) Einschätzungen einzuteilen. Die korrekten Einschätzer hätten demnach wie Durchschnittsregelanwender, die unkorrekten wie Additionsregelanwender und die inkorrekten ohne Vorstellung einer Mischungsregel geurteilt. Mit dieser Kategorisierung der Einschätzungen der Mischungsresultate konnte also ein direkter Vergleich mit den prädiktiven Regelkategorien angestellt werden. Dahinter stand die Frage, ob die Durchschnittsregelanwender aus Phase 1 beim *magic experiment* häufiger korrekt urteilten als die anderen Regelanwender. Hinsichtlich der Stabilitätshypothese interessierte insbesondere, welche Regelanwender mehr ihrer prädiktiven Regel entsprechend urteilten, d. h. welche Regeln sich im *magic experiment* als stabiler erwiesen.

Subjektive Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate

Mischungsart: Hat die Art des Mischungsresultates Einfluss auf die subjektive Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate? Hier bestand die Erwartung, dass sich die der Mischungsregel widersprechende Mischungsart verunsichernd und das erwartete Mischungsresultat bestätigend auf die subjektive Regelsicherheit auswirkte. Zudem wurde erwartet, dass die subjektive Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate mit zunehmendem Alter grösser wurde.

Intensitätskombination: Zeigte sich beim bewussten Umgang mit der prädiktiven Regel ein Unterschied zwischen den beiden Intensitätskombinationen? Da sich bei der Prädiktion im Umgang mit den beiden Intensitätskombinationen deutliche Unterschiede gezeigt hatten, werden die erhobenen Sicherheitsurteile nach Beobachtung der Mischungsvorgänge auch auf allfällige Unterschiede zwischen den beiden Intensitätskombinationen betrachtet.

Prädiktive Regelanwendung: Auch bei der Erfassung der subjektiven Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate war die zentrale Frage, ob sich die verschiedenen Regelanwender aus Phase 1 unterschieden hatten. Es wurde vermutet, dass das den prädiktiven Mischungsregeln entsprechende Resultat zu mehr Sicherheit führte als das den eigenen Regeln widersprechende. Konkret wurde erwartet, dass sich bei den Anwendern der Durchschnittsbildung aus Phase 1 das physikalisch korrekte Mischungsresultat und bei den Anwendern der Additionsregel das physikalisch unkorrekte Mischungsresultat auf eine hohe subjektive Sicherheit auswirkte. Umgekehrt konnte man vermuten, dass sich das physikalisch unkorrekte Mischungsresultat sich bei den Anwendern der Durchschnittsregel aus Phase 1 sowie das physikalisch korrekte Mischungsresultat bei den Anwendern der Additionsregel auf eine geringere subjektive Sicherheit auswirkte. Hinsichtlich eines Entwicklungsverlaufes bei dieser abhängigen Variablen konnte angenommen werden, dass die Durchschnittsregel mit zunehmendem Alter sicherer beibehalten und vertreten werden konnte, dies auch bei der physikalisch unkorrekten Mischungsart. Bei noch unsicheren Kindern oder Anwendern einer falschen Mischungsregel könnte die Regel aber auch infrage gestellt worden sein.

Da nicht ausgeschlossen werden konnte, dass ein generelles individuelles Sicherheitsniveau die Einschätzungen der subjektiven Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate beeinflusst haben könnte, wurde der Einfluss der subjektiven Sicherheit aus Phase 1 auf das Sicherheitsurteil in Phase 2 als Kovariate überprüft.

Die Einschätzungen der subjektiven Sicherheit der prädiktiven Regel nach Beobachtung der Mischungsresultate konnten zudem den drei Kategorien stabile (positive Werte), verunsicherte (Wert 0) und instabile (negative Werte) Regelanwender zugewiesen werden. Somit konnte der direkte Umgang mit ihrer prädiktiven Regel nach Beobachtung verschiedener Mischungsresultate beurteilt werden. Es wurden Unterschiede zwischen den verschiedenen Regelanwendern aus Phase 1 in der Bedingung des *magic experiment* erwartet: Die Anwender der korrekten Durchschnittsregel sollten beim *magic experiment* ihre Regel stabiler gegenüber dem gegenteiligen Beweis vertreten als die Anwender der falschen Additionsregel. Bei letzteren kann auch vermutet werden, dass die falsche Regel möglicherweise sogar zu Gunsten der korrekten Regel aufgegeben wurde.

3.3.2 Methode

Die Angaben zu Versuchspersonen, -material und -durchführung wurden bereits in Kapitel 3.1 dargestellt. An dieser Stelle wird nur noch das Versuchsdesign mit unabhängigen und abhängigen Variablen des *magic experiment* dargestellt.

Versuchsplan

Das *magic experiment* bezieht sich auf Phase 2 des gesamten Experimentes, d.h. auf die Erfassung der Reaktionen auf die physikalisch korrekten und unkorrekten Mischungen gleicher und unterschiedlicher Ursprungsfarbbintensitäten. Folgende Variablen lagen der Phase 2 des Experimentes zugrunde:

Unabhängige Variablen

- Faktor A (Messwiederholungsfaktor; zweistufig): Mischungsresultat bzw. -art
- Faktor B (Messwiederholungsfaktor; zweistufig): Farbbintensitätskombination
- Faktor C (Zwischensubjektfaktor; fünfstufig): Alter

- Faktor D (Zwischensubjektfaktor; zweistufig): Geschlecht

Der Versuchsplan entsprach einem *2x2-within-subjects-design* mit vier Versuchsdurchgängen pro Versuchsperson (vgl. Anhang C, S. 276ff.). Die Versuchsdurchgänge wurden in zwei verschiedenen Reihenfolgen dargeboten. Zu Beginn des Versuches wurde bei allen Versuchspersonen zuerst ein Probeversuchsdurchgang durchgeführt, bei welchem die Mischung von zwei farblosen Flüssigkeiten farbig (gleiches intensives Rotviolett wie bei den Versuchsdurchgängen) wurde. Die Angaben zum Alter und zur Geschlechtsverteilung der Versuchspersonen erfolgten bereits in Kapitel 3.1.

Abhängige Variablen

- Einschätzung der Mischungsresultate auf der Wahrnehmungsebene: Ist die Flüssigkeit durch Zugabe der anderen Flüssigkeit heller oder dunkler geworden oder gleich geblieben? Zur Vereinfachung der Aufgabe war auch ein Vergleich mit einem Referenzglas möglich. Die abhängige Variable lag mit drei Antwortkategorien auf einem nominalen Skalenniveau.
- Realitätseinschätzungen anhand der Zauberskala: Wie fest ist die Mischung gezaubert? Die Einschätzungen auf der Zauberskala waren von 0 bis 10 möglich, d. h. die abhängige Variable war proportionalskaliert (siehe Anhang D, S. 282).
- Diskrimination zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten: Diese abhängige Variable wurde berechnet aus der Differenz der Einschätzungen auf der Zauberskala beim physikalisch korrekten und physikalisch unkorrekten Mischungsresultat. Theoretisch waren Werte von –10 bis 10 möglich. Die Diskriminationswerte lagen intervallskaliert vor. Positive Werte entsprachen höheren Werten mit der Zauberskala bei den physikalisch unkorrekten Mischungen, d. h. korrekten Einschätzungen. Negative Werte entsprachen umgekehrt höheren Werten mit der Zauberskala bei den physikalisch korrekten Mischungen, d. h. sie entsprachen unkorrekten Einschätzungen. Zudem waren auch indifferente Einschätzungen möglich, wenn beide Mischungsarten auf der Zauberskala gleich hoch eingeschätzt wurden. Durch Kategorisierung in die drei Variablenstufen korrekte, indifferente und unkorrekte Einschätzungen lagen die Diskriminationen auch nominalskaliert vor.
- Subjektive Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenartigen Mischungsresultate: Wie sicher bist du, dass dein prädikti-

ves Urteil richtig war? Einschätzungen waren auf einer Rating-Skala von –5 bis 5 möglich (siehe Anhang D, S. 283). Die negativen Werte entsprachen einer Revidierung des prädiktiven Mischungsurteils: je sicherer die Revidierung war umso negativer wurde der Wert. Wenn nach der Rückmeldung absolute Unsicherheit im Hinblick auf das prädiktive Urteil bestand, wurde der Wert null angegeben. Zur statistischen Auswertung wurden die Daten in Werte von 0 bis 10 umkodiert. Die Daten der subjektiven Sicherheitseinschätzungen lagen intervallskaliert vor. Auch hier wurde eine Kategorisierung vorgenommen: Negative Einschätzungen entsprachen einer instabilen und positive Einschätzungen einer stabilen Regelanwendung. Werte von null wurden der Kategorie der unsicheren Regelanwendung zugeordnet.

3.3.3 Auswertung

Einschätzung der Mischungsergebnisse auf der Wahrnehmungsebene

Um bei den Einschätzungen der Mischungsergebnisse auf der Wahrnehmungsebene auszuschließen, dass sie möglicherweise durch Raten entstanden waren, wurden die Häufigkeiten der drei Antwortmöglichkeiten pro Versuchsdurchgang ausgezählt. Für jede der drei Antwortmöglichkeiten bestand eine Ratewahrscheinlichkeit von einem Drittel. Falls es überhaupt nötig wurde, die beobachteten Häufigkeiten gegen den Zufall statistisch abzusichern, wäre die Berechnung eines χ^2 -Tests mit der Erwartung von gleichen Werten bei allen drei Kategorien in Betracht gekommen. Aufgrund der vorliegenden Häufigkeiten korrekter und unkorrekter Wahrnehmungseinschätzungen der Mischungsergebnisse war auch die Überprüfung der Verteilungen zur Abhängigkeit von den verschiedenen Altersgruppen mit dem χ^2 -Test nach Pearson möglich. Zudem konnten die Verteilungen dieser Einschätzungskategorien auf die verschiedenen Regelanwendungskategorien aus Phase 1 auf einen Zusammenhang geprüft werden.

Realitätseinschätzungen anhand der Zauberskala

Varianzanalysen

Es wurde eine Varianzanalyse für die Einschätzungen auf der Zauberskala (Werte von 0 bis 10) der vier Versuchsdurchgänge mit den Messwiederholungsfaktoren Art des Mi-

schungsergebnisses und Farbintensitätskombination sowie den Zwischensubjektfaktoren Alter und Geschlecht durchgeführt. Es sollte geprüft werden, ob sich die Art des Mischungsergebnisses (physikalisch korrekte und unkorrekte Farbintensitätsmischungen) – alleine und in Abhängigkeit vom Alter – der Versuchspersonen auf die Realitätseinschätzung auswirkte. Zudem wurde überprüft, ob sich die Versuchspersonen bezüglich ihrer Einschätzungen des gezauberten Charakters der verschiedenen Mischungsergebnisse bei gleichen Intensitäten (Intensitätskombination 2 und 2) gegenüber derjenigen bei unterschiedlichen Intensitäten (Intensitätskombination 1 und 3) unterschieden. Ferner sollte abgeklärt werden, ob sich Unterschiede zwischen den Reihenfolgen der Versuchsdurchgänge (erneute Varianzanalyse mit diesem Faktor anstelle des Geschlechts) zeigten.

Um zu prüfen, ob sich die zu Regeln kategorisierten prädiktiven Mischungsurteile auf die Realitätseinschätzungen mit der Zauberskala auswirkten, wurden weitere Varianzanalysen berechnet. Da die Mischungsregeln nur innerhalb derselben Intensitätskombinationen bestimmt wurden, mussten die Varianzanalysen getrennt durchgeführt werden. Es wurde pro Intensitätskombination je eine Varianzanalyse mit dem Messwiederholungsfaktor Art des Mischungsergebnisses sowie den Zwischensubjektfaktoren prädiktive Regelanwendung, Alter und Geschlecht gerechnet. Auch hier wurden allfällige Effekte der Geschlechtszugehörigkeit und der Reihenfolge der Versuchsdurchgänge (erneute Varianzanalyse mit diesem Faktor anstelle des Geschlechts) überprüft.

Zur Überprüfung, ob die beiden Tricks (Wassertrick beim Probeversuchsdurchgang und Farbtrick bei den Versuchsdurchgängen) hinsichtlich der Einschätzung mit der Zauberskala unterschieden wurden, wurde eine Varianzanalyse mit dem Messwiederholungsfaktor Art des Tricks und dem Zwischensubjektfaktor Alter durchgeführt.

Im Anschluss an die Varianzanalysen wurden an bestimmten Stellen zur weiteren Aufklärung der Haupteffekte Post-Hoc-Tests nach Scheffé gerechnet. Bei auftretenden Interaktionen zwischen verschiedenen Faktoren wurden weitere Varianzanalysen, getrennt für die Stufen eines Faktors, berechnet.

Diskrimination zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten

Varianzanalysen

Es wurde eine Varianzanalyse für die berechnete Diskriminationsfähigkeit (Werte von -10 bis 10) bei beiden Intensitätskombinationen (Mischungen gleicher und unterschiedlicher Ursprungsintensitäten) mit dem Messwiederholungsfaktor Farbintensitätskombination und den Zwischensubjektfaktoren Alter und Geschlecht durchgeführt. Es sollte geprüft werden, ob sich die Diskriminationsfähigkeit zwischen den beiden Arten von Mischungsresultaten über die verschiedenen Altersstufen hinweg verändert. Es wurde zudem überprüft, ob sich die Versuchspersonen bezüglich ihrer Diskriminationsfähigkeit bei gleichen Intensitäten (Intensitätskombination 2 und 2) gegenüber derjenigen bei unterschiedlichen Intensitäten (Intensitätskombination 1 und 3) unterschieden. Weiter sollten Unterschiede zwischen den Geschlechtern überprüft werden.

Um zu prüfen, ob sich die zu Regeln kategorisierten prädiktiven Mischungsurteile auf die Diskriminierungsfähigkeit zwischen den beiden Mischungsarten auswirkten, wurden weitere Varianzanalysen berechnet. Es wurde getrennt für die beiden Intensitätskombinationen je eine zweifaktorielle univariate Varianzanalyse mit den Zwischensubjektfaktoren prädiktive Regelanwendung und Alter berechnet.

Im Anschluss an die Varianzanalysen wurden beim Auftreten von Haupteffekten Post-Hoc-Tests nach Scheffé gerechnet, um aussagen zu können, welche Stichproben bzw. Stufen eines Faktors sich unterschieden hatten. Bei auftretenden Wechselwirkungen zwischen verschiedenen Faktoren wurden weitere Varianzanalysen, getrennt für die einzelnen Stufen eines Faktors, berechnet.

χ^2 -Tests

Zum direkten Vergleich der Einschätzungen im *magic experiment* und den Einschätzungen in der Prädiktionsbedingungen aus Phase 1 wurden die Werte der Diskriminationen einer der drei Kategorien zugewiesen: Positive Diskriminierungen entsprachen korrekten Einschätzungen im *magic experiment*, keine Diskriminierung (Wert 0) entsprach einer indifferenten Einschätzung und negative Diskriminierungen entsprachen

unkorrekten Einschätzungen. Die Verteilungen der Versuchspersonen auf die beiden Variablen konnten in einer 3x3-Kreuztabelle dargestellt werden. Mit dem χ^2 -Test nach Pearson wurde überprüft, ob zwischen den beiden Variablen ein Zusammenhang bestand. Falls in zu vielen Zellen (mehr als 20% der Felder) zu geringe erwartete Häufigkeiten (< 5) ausgegeben wurden, wurde ein Vierfelder- χ^2 -Test mit Fishers exaktem Test für die relevantesten Kategorien durchgeführt (vgl. Brosius, 2002).

Subjektive Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsergebnisse

Varianzanalysen

Es wurde eine Varianzanalyse für die Einschätzungen der subjektiven Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsergebnisse (Wertebereich -5 bis 5 wurde in 0 bis 10 umkodiert) bei allen vier Versuchsdurchgängen mit den Messwiederholungsfaktoren Art des Mischungsergebnisses und Farbintensitätskombination sowie den Zwischensubjektfaktoren Alter und Geschlecht durchgeführt. Es sollte geprüft werden, ob sich die Beobachtung der verschiedenen Mischungsarten auf die subjektive Sicherheit ausgewirkt hatte. Zudem wurde überprüft, ob sich die Versuchspersonen bezüglich ihrer Regelsicherheit nach der Rückmeldung bei den beiden Intensitätskombinationen (gleiche und unterschiedliche) unterschieden. Ein allfälliger Effekt des Alters auf die subjektive Sicherheit nach Beobachtung der Mischungsergebnisse – alleine oder in Wechselwirkung mit einem der anderen Faktoren – wurde gleichzeitig mitgeprüft. Ferner war es auch bei dieser abhängigen Variablen von Interesse, abzuklären, ob die Geschlechtszugehörigkeit oder die Reihenfolge der Versuchsdurchgänge Auswirkungen auf die Einschätzungen hatten. Letzteres wurde mit einer erneuten Varianzanalyse nach der obigen Beschreibung und des Zwischensubjektfaktors Reihenfolge der Versuchsdurchgänge anstelle des Geschlechtsfaktors berechnet.

Um zu prüfen, ob sich die zu Regeln kategorisierten prädiktiven Mischungsurteile auf die die zugehörigen subjektiven Sicherheitsurteile nach Beobachtung des Mischungsvorganges auswirkten, wurden weitere Varianzanalysen berechnet. Da die Mischungsregeln nur innerhalb derselben Intensitätskombinationen bestimmt wurden, mussten die Varianzanalysen getrennt durchgeführt werden. Es wurde pro Intensitätskombination je

eine Varianzanalyse mit dem Messwiederholungsfaktor Art des Mischungsresultates sowie den Zwischensubjektfaktoren prädiktive Regelanwendung und Alter gerechnet.

Für alle Varianzanalysten galt: Bei auftretenden Interaktionen zwischen verschiedenen Faktoren wurden weitere Varianzanalysen, getrennt für die Stufen eines Faktors, berechnet. Zudem sollten zur weiteren Aufklärung der Haupteffekte im Anschluss an die Varianzanalyse noch Post-Hoc-Tests nach Scheffé gerechnet werden.

Da hier zusätzlich noch eine proportionalskalierte unabhängige Variable in Form der subjektiven Sicherheitseinschätzung der prädiktiven Regel aus Phase 1 vorlag, war es zudem möglich den Einfluss dieser Kovariate auf die spätere Einschätzung mit einer Kovarianzanalyse zu überprüfen. Dies wurde pro Intensitätskombination zusammen mit dem Messwiederholungsfaktor Art des Mischungsresultates sowie den Zwischensubjektfaktoren prädiktive Regelanwendung und Alter berechnet.

χ^2 -Tests

Zum direkten Vergleich der Einschätzungen im *magic experiment* und den Einschätzungen in der Prädiktionsbedingung aus Phase 1 wurden die Werte der subjektiven Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsresultate einer der drei Kategorien zugewiesen: stabile (positive Werte), unsichere (Wert 0) und instabile Regelanwendung (negative Werte). Bei der statistischen Auswertung mit der Kreuztabelle und dem χ^2 -Test wurde genau gleich wie bei der Diskrimination vorgegangen (vgl. S. 104).

3.3.4 Ergebnisse

Einschätzung der Mischungsresultate auf der Wahrnehmungsebene

Die Versuchspersonen hatten nach dem Mischungsvorgang jeweils zuerst einzuschätzen, ob sich die Ausgangsflüssigkeit des einen Glases durch die Zugabe der zweiten Flüssigkeit hinsichtlich der Farbhelligkeit verändert hatte. Die Helligkeit der zu beurteilenden ursprünglichen Ausgangsflüssigkeit blieb auch nach dem Mischungsvorgang durch das Referenzglas, deren Flüssigkeit gleich farbintensiv war wie diese Ausgangsflüssigkeit, präsent. So konnte also auch einfach Mischungsresultat und Flüssigkeit im

Referenzglas, welche auch beide das gleiche Volumen hatten, miteinander verglichen werden. Die Versuchspersonen hatten jeweils drei Antwortkategorien zur Auswahl, wovon nur eine davon die richtige war. Es zeigten sich bei allen vier Versuchsdurchgängen Fehler bei der Einschätzung der Mischungsresultate auf der Wahrnehmungsebene. Dennoch lagen die insgesamt doch geringen Fehlerquoten bei allen vier Versuchsdurchgängen deutlich unter der Ratewahrscheinlichkeit von einem Drittel pro Antwortkategorie.

Folgende Wahrnehmungsfehler wurden bei den einzelnen Versuchsdurchgängen beobachtet:

- *Versuchsdurchgang „gleiche Intensität, physikalisch korrektes Mischungsresultat“ (gr):* Das Mischungsresultat blieb genau gleich hell wie das Vergleichsglas. Insgesamt beurteilten aber 21 Versuchspersonen dieses Resultat anders, nämlich 15 antworteten, es sei heller geworden, und 6, es sei dunkler geworden anstatt gleich geblieben. Dieser Versuchsdurchgang hatte mit Abstand am meisten Schwierigkeiten bei der Einschätzung des Mischungsresultates auf der Wahrnehmungsebene geboten.
- *Versuchsdurchgang „gleiche Intensität, physikalisch unkorrektes Mischungsresultat“ (gf):* Hier war das Mischungsresultat dunkler geworden, was 4 Versuchspersonen anders einschätzten. Drei der falsche Einschätzer meinten, die Helligkeit sei gleich geblieben, und eine Person wählte sogar die Kategorie „ist heller anstatt dunkler geworden“.
- *Versuchsdurchgang „unterschiedliche Intensität, physikalisch korrektes Mischungsresultat“ (ur):* Da die hellere in die dunklere Ausgangsflüssigkeit geleert wurde, war das Mischungsresultat heller als die Referenzkomponente. Insgesamt 8 Versuchspersonen schätzten dies falsch ein. 7 urteilten, dass es gleich geblieben sei, und 1 Person meinte sogar, dass es dunkler anstatt heller geworden sei.
- *Versuchsdurchgang „unterschiedliche Intensität, physikalisch unkorrektes Mischungsresultat“ (uf):* Hier war das Mischungsresultat dunkler als die Referenzflüssigkeit. 4 Versuchspersonen schätzten dies falsch ein: 3 bei „ist geblieben anstatt dunkler geworden“ und 1 bei „ist heller geworden“.

Besonders auffällig erscheint hier, dass bei den physikalisch korrekten Mischungsresultaten viel mehr Fehler entstanden sind als bei den unkorrekten. Während bei unkorrekten Mischungsresultaten nur 8 Fehler gemacht wurden, waren es bei den korrekten 29.

Die Wahrnehmungsfehler bei der Einschätzung der Mischungsresultate waren im Altersverlauf unterschiedlich häufig. Abbildung 3.3.1 zeigt die Fehler pro Altersgruppe in Abhängigkeit von der Anzahl Fehler.

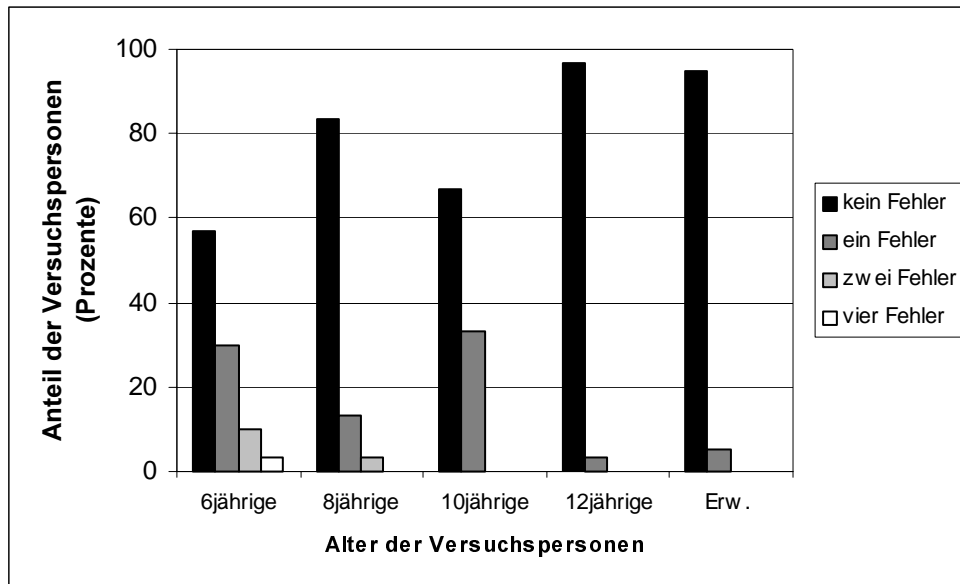


Abbildung 3.3.1. Prozentualer Anteil der Versuchspersonen in Abhängigkeit von der Anzahl Fehler bei der Wahrnehmungseinschätzung der Mischungsresultate pro Altersgruppe.

Der Abbildung lässt sich entnehmen, dass die Sechsjährigen am meisten Fehler bei der Wahrnehmungseinschätzung machten. Die Häufigkeiten der richtigen Einschätzer (kein Fehler) nahmen aber nicht mit zunehmendem Altersverlauf stetig zu, da die Zehnjährigen mehr Fehler machten als die Achtjährigen. Mehrfach falsche Einschätzer waren aber nur bei den beiden jüngsten Gruppen zu verzeichnen. Allerdings schätzte keine einzige Versuchsperson drei der vier Mischungsresultate falsch ein und nur eine Person schätzte alle vier falsch ein. Dabei handelt es sich um einen Sechsjährigen, bei welchem schon während der Versuchsdurchführung klar war, dass er hell und dunkel verwechselte. Auf diesen Knaben gingen auch die drei „Ausreisser“ bei den Versuchsdurchgängen *gf*, *ur* und *uf* zurück (siehe oben), welche dadurch auffielen, dass sie nicht nur um eine Stufe daneben lagen (z. B. gleich geblieben anstatt dunkler geworden), sondern zwei Stufen von der richtigen Helligkeit abwichen (z. B. dunkler anstatt heller geworden). Hätte dieser Sechsjährige, die Begriffe „heller,“ und „dunkler“ korrekt verwendet, hätte er nur eine falsche Wahrnehmungseinschätzung (bei *gr*) gemacht. Aufgrund des Vortests konnte auch vorausgesetzt werden, dass er ein Verständnis für Seriation hatte und die

Wahrnehmungsdiskrepanzen unterscheiden konnte. Zwei Fehler waren (dieser Sechsjährige ausgenommen) das höchste und dies nur bei den Sechs- und Achtjährigen.

Die Überprüfung, ob sich die Anzahl korrekter (kein Fehler) und unkorrekter (ein bis vier Fehler) Wahrnehmungseinschätzungen über den Altersverlauf hin unterschieden, wurde mit dem χ^2 -Test nach Pearson durchgeführt. Es zeigte sich, dass sich die Verteilungen bei den verschiedenen Altersgruppen höchst signifikant unterschieden, $\chi^2(4, N = 140) = 20.52; p < .001$. Nebst den bereits erwähnten vielen unkorrekten Einschätzungen bei den Sechsjährigen führten auch die vielen korrekten Einschätzungen bei den Zwölfjährigen zu diesem Resultat.

Zudem wurde anhand der Kreuztabelle überprüft, ob die falschen Einschätzer der Mischungsresultate auf der Wahrnehmungsebene auch falsche Regelanwender waren. Es zeigte sich aber sowohl bei der gleichen als auch der unterschiedlichen Intensitätskombination, dass kein Zusammenhang zwischen der Art der Prädiktion des Mischungsresultates und der Einschätzung dieses Resultates nach der Mischung auf der Wahrnehmungsebene bestand. Eine mögliche Wahrnehmungsverzerrung durch eine vorhandene individuelle Mischungsregel lag demnach nicht vor.

Realitätseinschätzungen anhand der Zauberskala

Die Realitätseinschätzungen der vier verschiedenen Mischungsresultate erfolgten anhand der Zauberskala (Werte 0 bis 10), wobei der Wert null einem nicht-gezauberten bzw. nicht-getricksten und der Wert zehn einem maximal gezauberten bzw. getricksten Mischungsresultat entsprach.

Die Abbildung 3.3.2 stellt die Mittelwerte für jeden der vier Versuchsdurchgänge in Abhängigkeit von der Altersgruppe dar.

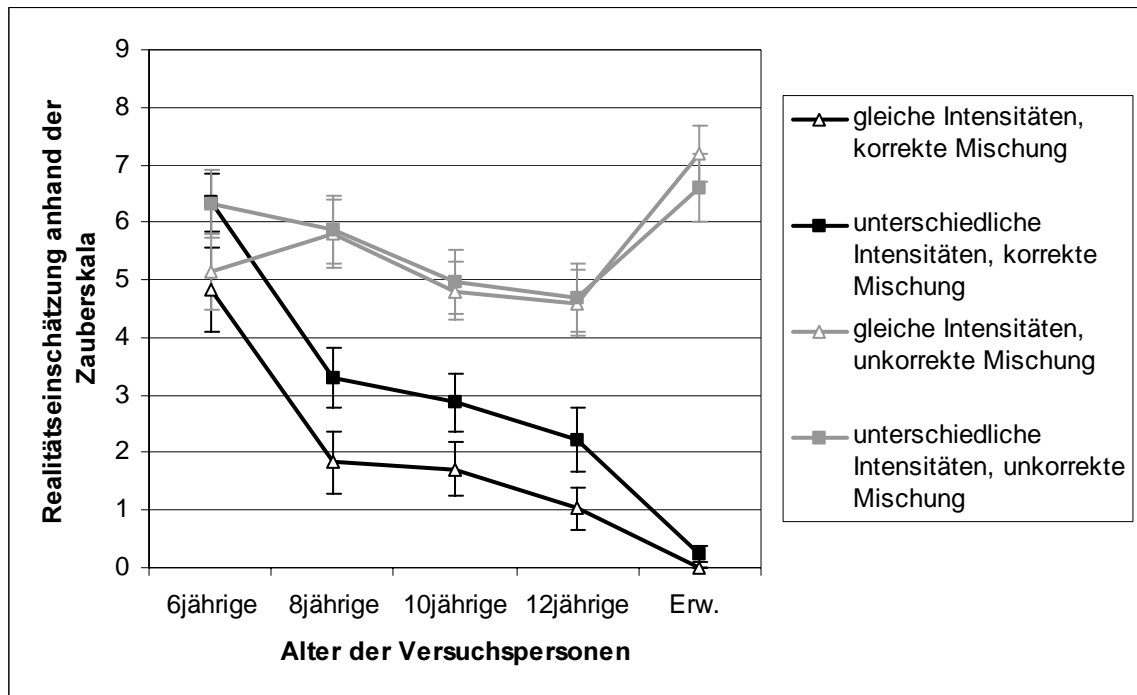


Abbildung 3.3.2. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) in Abhängigkeit von der Intensitätskombination, Art des Mischungsergebnisses und von der Altersgruppe.

Zur Überprüfung der Effekte der zwei Messwiederholungsfaktoren Farbintensitätskombination und Art des Mischungsergebnisses sowie den beiden Zwischensubjektfaktoren Alter und Geschlecht auf die Einschätzungen mit der Zauberskala wurde eine vierfaktorielle Varianzanalyse nach dem Allgemeinen linearen Modell berechnet. Es zeigten sich höchst signifikante Haupteffekte der Faktoren Farbintensitätskombination ($F(1, 130) = 12.94, p < .001$), Art des Mischungsergebnisses ($F(1, 130) = 165.81, p < .001$) und Alter ($F(4, 130) = 7.18, p < .001$). Ausserdem zeigten sich signifikante Wechselwirkungen zwischen Farbintensitätskombination und Art des Mischungsergebnisses ($F(1, 130) = 6.37, p < .05$) sowie zwischen Art des Mischungsergebnisses und Alter ($F(4, 130) = 16.2, p < .001$). Die Geschlechtszugehörigkeit hatte keinen differenziellen Einfluss auf die Urteile. Eine erneute Durchführung der Varianzanalyse mit dem Faktor Reihenfolge der Versuchsdurchgänge anstelle des Geschlechtseffektors zeigte, dass die Reihenfolge weder einen direkten Haupteffekt auf die Einschätzungen mit der Zauberskala hatte noch in einer Wechselwirkung mit einem anderen Faktor statistische Bedeutung erlangte.

Bedingt durch das Auftreten der Wechselwirkung zwischen den Faktoren Art des Mischungsergebnisses und Alter wurden, getrennt für die einzelnen Altersgruppen, Varianzanalysen mit den Messwiederholungsfaktoren Art des Mischungsergebnisses und Farbin-

tensitätskombination berechnet. In Abbildung 3.3.2 ist zu sehen, dass die Unterschiede zwischen den physikalisch korrekten und unkorrekten Mischungen mit zunehmendem Alter immer stärker berücksichtigt wurden. Während bei den Sechsjährigen der Unterschied nicht signifikant war, zeigten sich bei allen Altersgruppen höchst signifikante Unterscheidungen zwischen den beiden Arten von Mischungsergebnissen: Achtjährige: $F(1, 29) = 55.14, p < .001$; Zehnjährige: $F(1, 29) = 24.97, p < .001$; Zwölfjährige: $F(1, 29) = 19.25, p < .001$; Erwachsene: $F(1, 19) = 191.17, p < .001$.

Bedingt durch das Auftreten der Wechselwirkung zwischen den Faktoren Farbintensitätskombination und Art des Mischungsergebnisses wurden, getrennt für die beiden Arten des Mischungsergebnisses, Varianzanalysen mit dem Faktor Farbintensitätskombination berechnet. Aus obiger Abbildung wird deutlich, dass die beiden Intensitätskombinationen bei den physikalisch korrekten Mischungsergebnissen (dunklere Linien in Abbildung 3.3.2) auf der Zauberskala verschieden eingeschätzt wurden: die physikalisch korrekte Mischung von zwei unterschiedlich farbintensiven Flüssigkeiten wurde als mehr gezaubert eingeschätzt als die Mischung von zwei gleich farbintensiven Flüssigkeiten. Die Varianzanalyse errechnete denn auch einen höchst signifikanten Haupteffekt der Farbintensitätskombination auf die Einschätzung mit der Zauberskala bei den physikalisch korrekten Mischungsergebnissen, $F(1, 139) = 24.23, p < .001$. Bei den physikalisch unkorrekten Mischungsergebnissen zeigte die Farbintensitätskombination hingegen keinen statistisch bedeutsamen Effekt.

Realitätseinschätzungen in Abhängigkeit von der prädiktiven Regelanwendung

Weil die Einschätzungen zu konsistenten Regelanwendungen nur innerhalb derselben Intensitätskombination zugeordnet wurden, mussten die Berechnungen getrennt für jede Intensitätskombination durchgeführt werden.

Gleiche Intensitätskombination

Zur Überprüfung der Effekte des Messwiederholungsfaktors Art des Mischungsergebnisses sowie den drei Zwischensubjektfaktoren konsistente prädiktive Regelanwendung (ohne die vier Anwender der Subtraktionsregel), Alter und Geschlecht auf die Einschätzungen mit der Zauberskala wurde eine vierfaktorielle Varianzanalyse nach dem Allgemeinen linearen Modell berechnet. Es zeigten sich bei der gleichen Intensitätskombi-

nation signifikante Haupteffekte der Faktoren Art des Mischungsresultates ($F(1, 111) = 62.56, p < .001$) und Alter ($F(4, 111) = 3.21, p < .05$). Ausserdem resultierten signifikante Wechselwirkungen zwischen Art des Mischungsresultates und Alter ($F(4, 111) = 6.04, p < .001$) sowie zwischen Art des Mischungsresultates und prädiktive Regelanwendung ($F(2, 111) = 4.78, p = .01$). Die Geschlechtszugehörigkeit hatte keinen differenziellen Einfluss auf die Urteile. Eine erneute Durchführung der Varianzanalyse mit dem Faktor Reihenfolge der Versuchsdurchgänge anstelle des Geschlechtstfaktors zeigte, dass die Reihenfolge weder einen direkten Haupteffekt auf die Einschätzungen mit der Zauberskala hatte noch in einer Wechselwirkung mit einem anderen Faktor statistische Bedeutung erlangte.

Bedingt durch das Auftreten der Wechselwirkung zwischen den Faktoren Art des Mischungsresultates und prädiktive Mischungsregel wurden Varianzanalysen, getrennt für die einzelnen Gruppen von Regelanwendern, mit den Messwiederholungsfaktor Art des Mischungsresultates berechnet. Es zeigten sich höchst signifikante Unterscheidungen der Mischungsresultate bei den Durchschnittsbildern ($F(1, 74) = 125.58, p < .001$) und den inkonsistenten Regelanwendern ($F(1, 33) = 12.37, p = .001$). Den Addierern gelang die Unterscheidung der beiden Mischungsresultate nicht in statistisch relevantem Ausmass.

Die Abbildung 3.3.3 zeigt die Mittelwerte und Standardfehler für die beiden Arten von Mischungsresultaten in Abhängigkeit von der konsistenten prädiktiven Mischungsregel. Dabei zeigt sich deutlich, wie die verschiedenen Regelanwender die beiden Mischungsresultate auf der Zauberskala unterschieden und insbesondere, dass das Ausmass dieser Unterscheidung von der Art der prädiktiven Regelanwendung abhing.

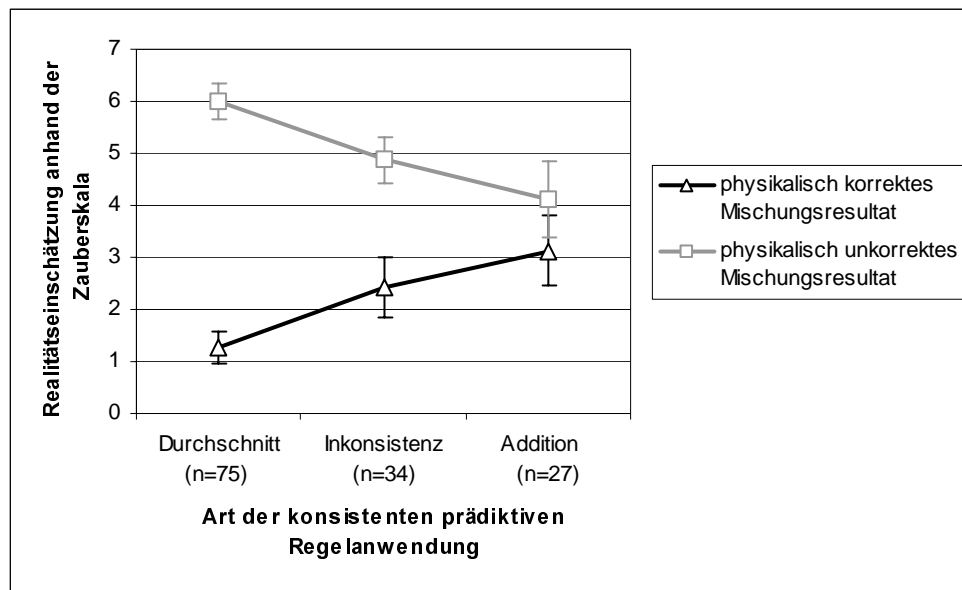


Abbildung 3.3.3. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) in Abhängigkeit von der konsistenten prädiktiven Regelanwendung und der Art des Mischungsresultates bei der gleichen Intensitätskombination.

Bedingt durch die Wechselwirkung zwischen der Art des Mischungsresultates und des Alters wurden zwei getrennte univariate Varianzanalysen für jede Mischungsart mit dem Faktor Alter gerechnet. Es zeigte sich, dass der Alterseffekt bei den unkorrekten Mischungsresultaten ($F(4, 140) = 2.76, p < .05$) bei weitem nicht so deutlich war wie bei den korrekten Mischungen ($F(4, 140) = 11.18, p < .001$).

Entwicklungsverläufe bei den verschiedenen Regelanwendern

Bei der physikalisch unkorrekten Mischungsart gab es bei allen drei Regelanwendern keinen signifikanten Effekt des Alters. Abbildung 3.3.4 ist zu entnehmen, dass deutliche Unterschiede im Altersverlauf nur bei den Additionsregelanwendern vorlagen. Hier zeigte sich aber kein Alterseffekt, weil die Stichprobe mit durchschnittlich sieben Versuchspersonen pro Altersgruppe zu klein war.

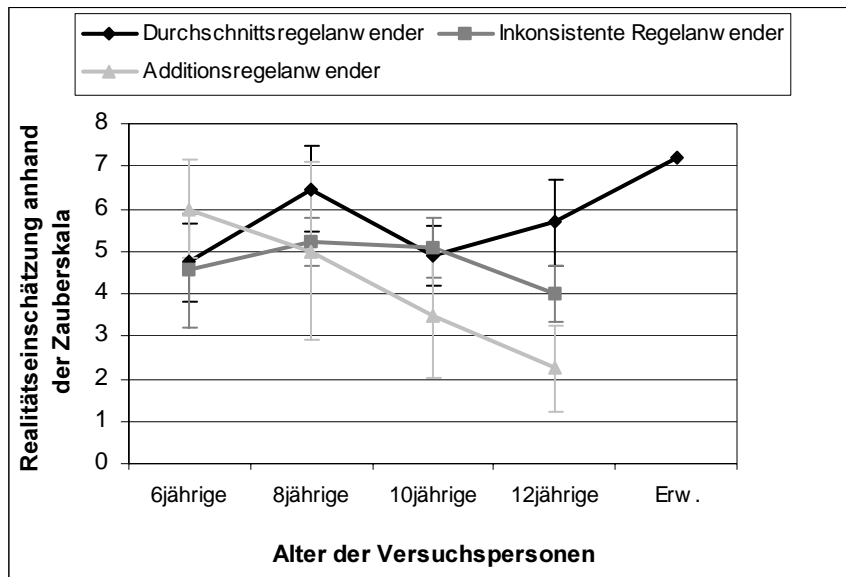


Abbildung 3.3.4 . Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) bei der physikalisch unkorrekten Mischungsart der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe.

Bei der physikalisch korrekten Mischungsart (Abbildung 3.3.5) hingegen war der Altersverlauf bei den Durchschnittsregelanwendern höchst signifikant, $F(4, 74) = 7.5$, $p < .001$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich alle Altersgruppen von den Sechsjährigen bei der Einschätzung auf der Zauberskala signifikant unterschieden (alle mit $p < .01$ ausser beim Vergleich Erwachsene und Sechsjährige mit $p < .001$). Auch bei den Additionsregelanwendern war der Effekt des Alters sehr signifikant, $F(3, 27) = 4.8$, $p = .01$. Hier verlief der Entwicklungsverlauf aber nicht stetig abnehmend, da die Zehnjährigen das physikalisch korrekte Mischungsresultat als mehr gezaubert eingeschätzt hatten, als die Achtjährigen. Der Post-Hoc-Test nach Scheffé ergab, dass sich die Sechsjährigen und die Zwölfjährigen bei der Einschätzung auf der Zauberskala signifikant unterschieden ($p < .05$), bei den Sechs- und den Achtjährigen zeigte sich ein Trend ($p = .053$).

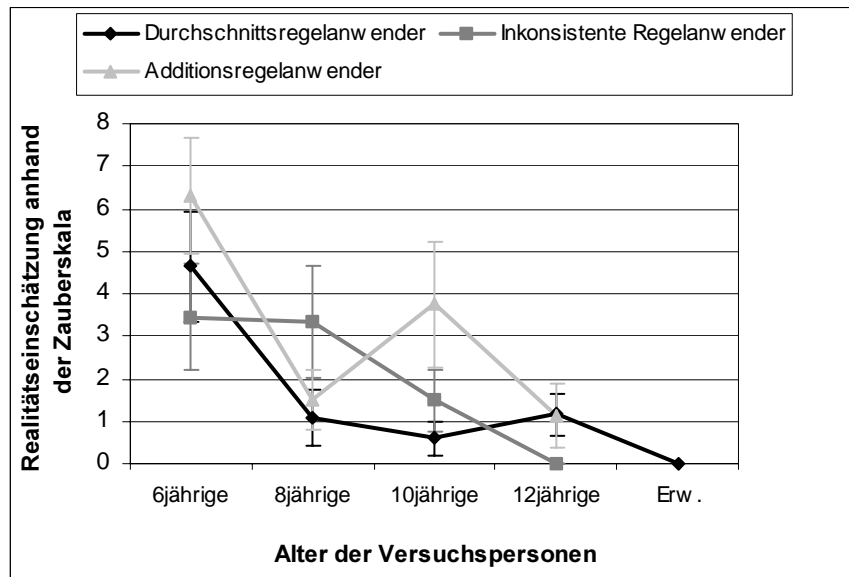


Abbildung 3.3.5. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) bei der physikalisch korrekten Mischungsart der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe.

Unterschiedliche Intensitätskombination

Zur Überprüfung der Effekte des Messwiederholungsfaktors Art des Mischungsergebnisses sowie den drei Zwischensubjektfaktoren konsistente prädiktive Regelanwendung (ohne die einzige Anwenderin der Zentrierungsregel), Alter und Geschlecht auf die Einschätzungen mit der Zauberskala wurde eine vierfaktorielle Varianzanalyse nach dem Allgemeinen linearen Modell berechnet. Es zeigten sich auch bei der unterschiedlichen Intensitätskombination signifikante Haupteffekte der Faktoren Art des Mischungsergebnisses ($F(1, 112) = 16.57, p < .001$) und Alter ($F(4, 112) = 7.06, p < .001$). Ausserdem zeigten sich – wie schon bei der gleichen Intensitätskombination – signifikante Wechselwirkungen zwischen Art des Mischungsergebnisses und Alter ($F(4, 112) = 5.39, p = .001$) sowie zwischen Art des Mischungsergebnisses und prädiktive Regelanwendung ($F(2, 112) = 4.28, p < .05$). Die Geschlechtszugehörigkeit hatte keinen differenziellen Einfluss auf die Urteile. Eine erneute Durchführung der Varianzanalyse mit dem Faktor Reihenfolge der Versuchsdurchgänge anstelle des Geschlechtseffektors zeigte, dass die Reihenfolge weder einen direkten Haupteffekt auf die Einschätzungen mit der Zauberskala hatte noch in einer Wechselwirkung mit einem anderen Faktor statistische Bedeutung erlangte.

Bedingt durch das Auftreten der Wechselwirkung zwischen den Faktoren Art des Mischungsresultates und prädiktive Mischungsregel wurden, getrennt für die einzelnen Gruppen von Regelanwendern, Varianzanalysen mit dem Messwiederholungsfaktor Art des Mischungsresultates berechnet. Die Varianzanalyse ergab einen höchst signifikanten Haupteffekt der Art des Mischungsresultates auf die Realitätseinschätzung bei den Durchschnittsbildern, $F(1, 87) = 83.32, p < .001$. Die beiden anderen Gruppen von Regelanwendern unterschieden die beiden Resultate nicht in statistisch relevantem Ausmass. Die Abbildung 3.3.6 zeigt die Mittelwerte und Standardfehler für die beiden Arten von Mischungsresultaten in Abhängigkeit von der konsistenten prädiktiven Mischungsregel bei der unterschiedlichen Intensitätskombination. Es zeigt sich – im Unterschied zur gleichen Intensitätskombination –, dass die inkonsistenten Regelanwender die beiden Mischungsresultate noch geringer unterschieden als die Addierer. Bei beiden Intensitätskombinationen waren aber die Durchschnittsbilder die deutlichsten Diskriminierer.

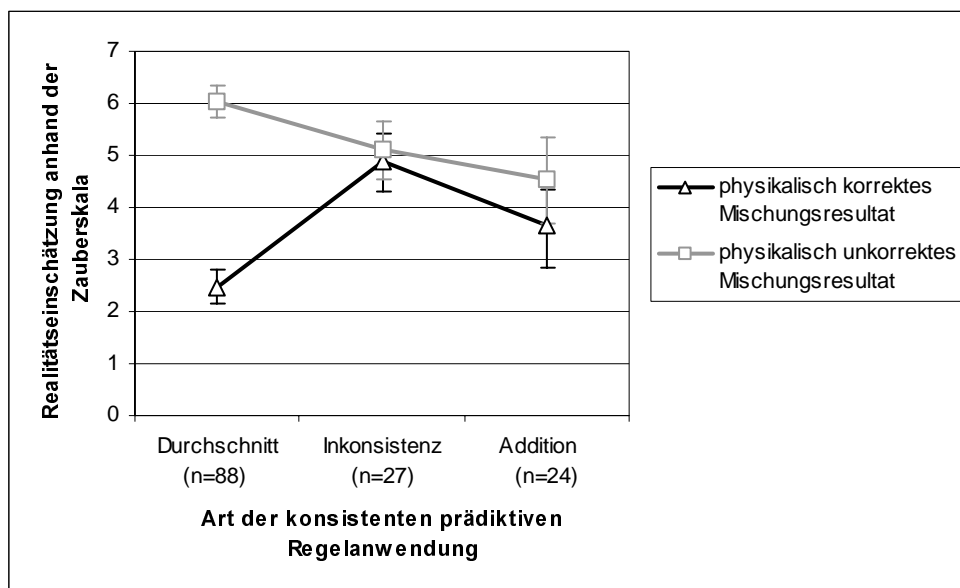


Abbildung 3.3.6. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) in Abhängigkeit von der konsistenten prädiktiven Regelanwendung und der Art des Mischungsresultates bei der unterschiedlichen Intensitätskombination.

Bedingt durch die Wechselwirkung zwischen der Art des Mischungsresultates und des Alters wurden zwei getrennte univariate Varianzanalysen für jede Mischungsart mit dem Faktor Alter gerechnet. Es zeigte sich, dass es bei den unkorrekten Mischungsresultaten gar keinen Alterseffekt, bei den korrekten Mischungen hingegen einen höchst signifikanten Alterseffekt gab ($F(4, 140) = 18.53, p < .001$).

Entwicklungsverläufe bei den verschiedenen Regelanwendern

Bei der physikalisch unkorrekten Mischungsart (Abbildung 3.3.7) gab es bei allen drei Regelanwendern keinen signifikanten Effekt des Alters. Beim Verlauf der Additions- und inkonsistenten Regelanwender wird aber angenommen, dass sich ein statistisch signifikanter Altersverlauf gezeigt hätte, wenn die Stichproben grösser gewesen wären als bei den vorliegenden mit durchschnittlich fünf bis sieben Versuchspersonen pro Altersgruppe.

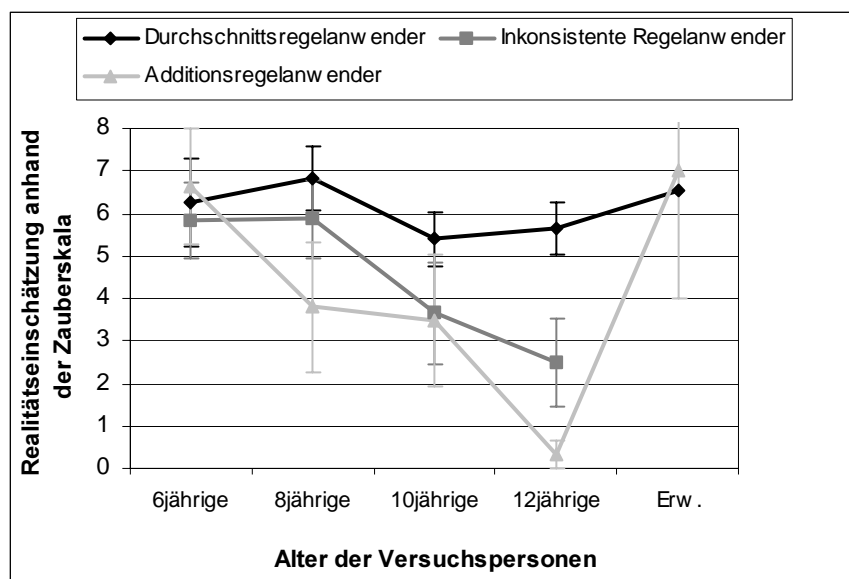


Abbildung 3.3.7. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) bei der physikalisch unkorrekten Mischungsart der unterschiedlichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe.

Bei der physikalisch korrekten Mischungsart (Abbildung 3.3.8) zeigte sich – wie schon bei der gleichen Intensitätskombination – ein höchst signifikanter Effekt des Alters bei den Durchschnittsregelanwendern, $F(4, 87) = 10.39$, $p < .001$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich alle Altersgruppen von den Sechsjährigen (Erwachsene und Sechsjährige, $p < .001$, Zwölf- und Sechsjährige, $p = .01$, Zehn- und Sechsjährige, $p < .01$, Acht- und Sechsjährige, $p = .053$) sowie die Erwachsenen und die Achtjährigen ($p < .05$) bei der Realitätseinschätzung signifikant unterschieden. Bei den Additionsregelanwendern ergab die Überprüfung des Alterseffekts einen Trend, $F(4, 24) = 2.39$, $p = .087$.

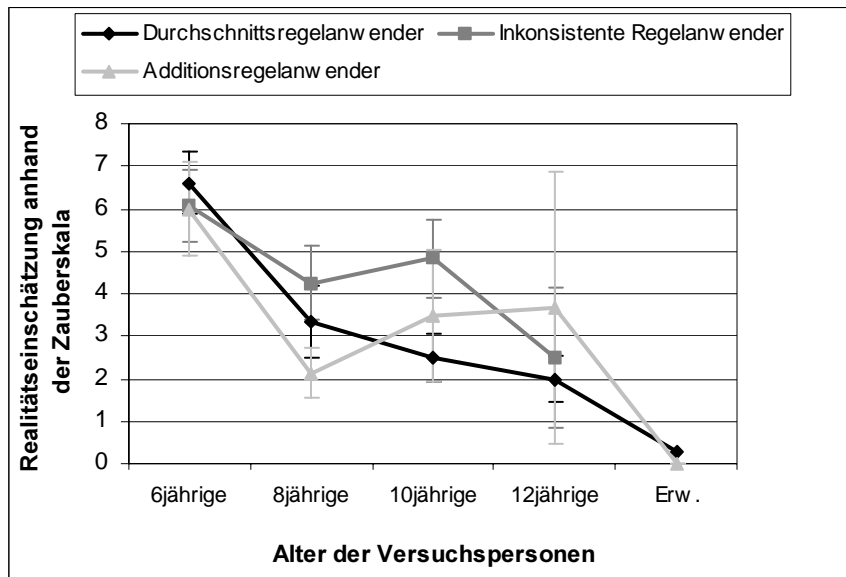


Abbildung 3.3.8. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) bei der physikalisch korrekten Mischungsart der unterschiedlichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe.

Zauberskala beim Probeversuchsdurchgang und Vergleich Probe- und Versuchsdurchgänge

Die Berechnung einer univariaten Varianzanalyse mit der abhängigen Variablen Einschätzung auf der Zauberskala beim Probeversuchsdurchgang und den unabhängigen Variablen Alter und prädiktive Mischungsregel zeigte nur einen Alterseffekt, $F(4, 140) = 3.29$, $p < .05$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich die Zehnjährigen und Erwachsenen sowie die Zwölfjährigen und Erwachsenen signifikant bei der Einschätzung auf der Zauberskala unterschieden (beide mit $p < .05$). Der Verlauf der Einschätzung des Mischungsergebnisses beim Probeversuchsdurchgang (Wassertrick) über die verschiedenen Altersgruppen ist in Abbildung 3.3.9 (obere schwarze Linie) dargestellt.

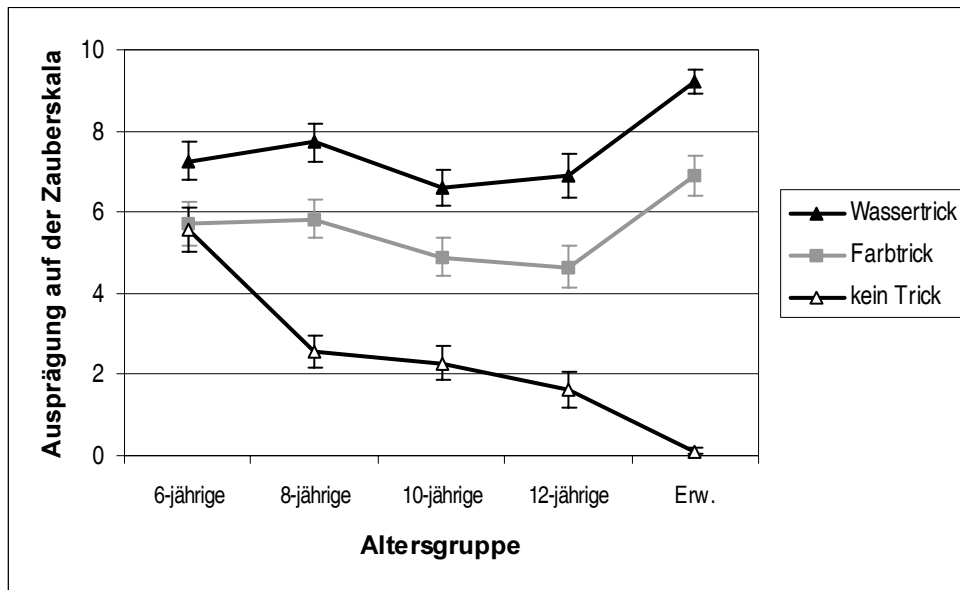


Abbildung 3.3.9. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala (0-10) beim Probeversuchsdurchgang (Wassertrick) sowie bei den physikalisch korrekten (kein Trick) und unkorrekten (Farbtrick) Mischungsresultaten in Abhängigkeit von der Altersgruppe.

Die obige Abbildung zeigt, wie unterschiedlich die verschiedenen Mischungsresultate beim Probeversuchsdurchgang (Wassertrick), bei den physikalisch korrekten (kein Trick; Mittelwerte der beiden Versuchsdurchgänge gr und ur) und den physikalisch unkorrekten Mischungsresultaten (Farbtrick; Mittelwerte der beiden Versuchsdurchgänge gf und uf) auf der Zauberskala eingeschätzt wurden. Es zeigte sich, dass die beiden Tricks im Altersverlauf ähnlich eingeschätzt wurden, in der Höhe der Einschätzungen aber ein deutlicher Unterschied bestand. Da der Unterschied zwischen dem Farbtrick und den physikalisch möglichen Mischungsresultaten (kein Trick) schon vorher berechnet wurde (Messwiederholungsfaktor: Art des Mischungsresultates; siehe S. 109) und damals schon höchst signifikant war, ist durch Berechnung des Unterschiedes zwischen Wassertrick und „kein Trick“ nur noch eine Steigerung in der Höhe des F-Wertes zu erwarten, weshalb auf diese Berechnung verzichtet wurde. Es wurde aber eine Varianzanalyse nach dem Allgemeinen Linearen Modell mit dem Messwiederholungsfaktor „verschiedene Tricks“ durchgeführt. Dabei zeigten sich höchst signifikante Haupteffekte von „verschiedene Tricks“ ($F(1, 135) = 55.58, p < .001$) und Alter ($F(4, 135) = 4.68, p = .001$). Der Post-Hoc-Test nach Scheffé zu den Altersunterschieden ergab, dass sich die Zehn- und Zwölfjährigen je von den Erwachsenen unterschieden ($p < .001$ bzw. $p = .001$), d. h. die älteren Kinder schätzten den Wassertrick als weniger gezaubert ein als die jüngeren Kinder. Interessant war zudem, dass auch die Sechsjährigen die beiden

Tricks bei der Einschätzung auf der Zauberskala unterschieden ($F(1, 29) = 6.52, p < .05$).

Diskrimination zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten

Die Diskriminationsfähigkeit wurde aus der Differenz zwischen Realitätseinschätzungen bei physikalisch unkorrekten und korrekten Mischungsresultaten pro Intensitätskombination berechnet.

Generelle Diskriminierungsleistungen

Um den Einfluss des Alters und der Intensitätskombination auf die Diskriminierungsleistungen mit der Zauberskala zu überprüfen, wurde eine Varianzanalyse nach dem Allgemeinen Linearen Modell mit den Messwiederholungsfaktor Farbintensitätskombination und den Zwischensubjektfaktoren Alter sowie Geschlecht durchgeführt. Es ergaben sich signifikante Haupteffekte der Faktoren Alter ($F(4, 130) = 16.2, p < .001$) und Farbintensitätskombination ($F(1,130) = 6.37, p < .05$). Die Geschlechtszugehörigkeit hatte keinen differenziellen Einfluss auf die Diskriminierungseinschätzungen mit der Zauberskala. Der Post-Hoc-Test nach Scheffé zur Berechnung von Altersunterschieden für beide Intensitätskombinationen zusammen zeigte signifikante Differenzen zwischen den Erwachsenen und allen Altersgruppen der Kinder: mit Sechsjährigen: $p < .001$; mit Achtjährigen: $p < .01$; mit Zehnjährigen: $p < .001$; mit Zwölfjährigen: $p = .001$. Aber auch zwischen den Sechsjährigen und allen älteren Kindern waren signifikante Unterschiede zu beobachten: mit Achtjährigen: $p < .01$; mit Zehnjährigen: $p < .05$; mit Zwölfjährigen: $p < .01$.

Abbildung 3.3.10 veranschaulicht die beiden statistisch bedeutsamen Effekte von Alter und Intensitätskombination auf die Diskriminierungsfähigkeit.

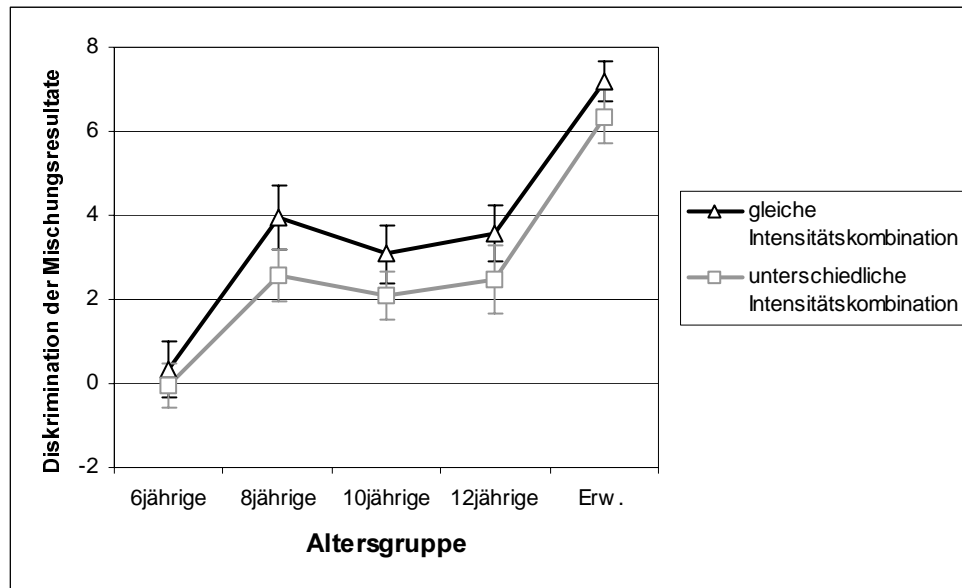


Abbildung 3.3.10. Mittelwerte und Standardfehler der Diskriminierungen zwischen den physikalisch unkorrekten und korrekten Mischungsergebnissen bei den beiden Intensitätskombinationen in Abhängigkeit von der Altersgruppe. Die Ausprägungen der Diskriminierungen sind definiert durch die Differenz der Einschätzungen mit der Zauberskala zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen, d. h. positive Diskriminierungen entsprechen höheren Werten mit der Zauberskala bei den physikalisch unkorrekten Mischungen als bei den korrekten Mischungen und somit korrekten Realitätseinschätzungen.

Die Diskriminierungswerte wurden zur weiteren Auswertung den folgenden drei Kategorien zugewiesen: physikalisch korrekte, indifferente und unkorrekte Einschätzung der beiden Mischungsergebnisse. In Abbildung 3.3.11 sind die prozentualen Anteile dieser drei Kategorien pro Altersgruppe für die Mischungsergebnisse der gleichen Intensitätskombination dargestellt.

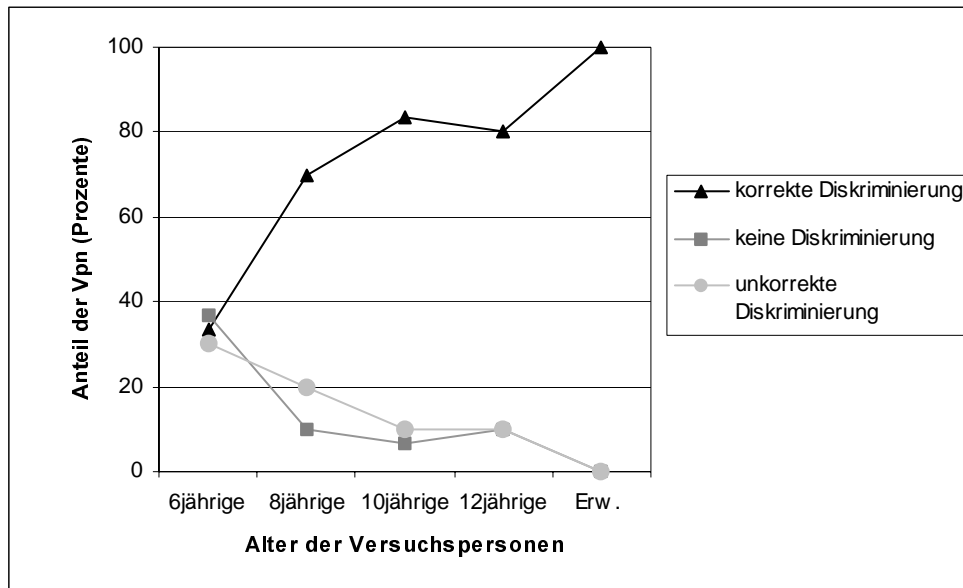


Abbildung 3.3.11. Korrekte, indifferente und unkorrekte Diskriminierungen zwischen den beiden Mischungsarten bei der gleichen Intensitätskombination in Prozenten pro Altersgruppe.

Die Überprüfung auf Unabhängigkeit der Verteilungen unterschiedlicher Diskriminierungen auf die verschiedenen Altersgruppen mit dem χ^2 -Test nach Pearson war nicht möglich, da die Voraussetzungen der Zellenbesetzung bezüglich der erwarteten Häufigkeiten < 5 nicht erfüllt waren. Es zeigt sich aber auch auf der deskriptiven Ebene, dass die Häufigkeiten korrekter Diskriminierungen sehr deutlich mit zunehmendem Alter der Versuchspersonen anstiegen und sie sich von den anderen Diskriminierungsarten deutlich unterschieden.

Die prozentualen Anteile der Diskriminierungsarten bei der unterschiedlichen Intensitätskombination sind in Abbildung 3.3.12 in Form eines Liniendiagramms dargestellt.

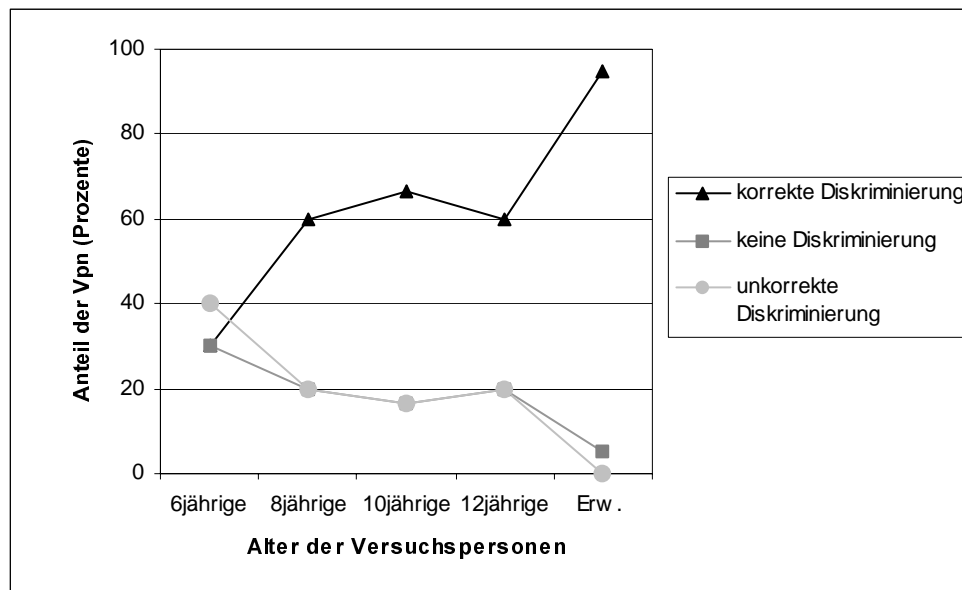


Abbildung 3.3.12. Korrekte, indifferente und unkorrekte Diskriminierungen zwischen den beiden Mischungsarten bei der unterschiedlichen Intensitätskombination in Prozenten pro Altersgruppe.

Die Überprüfung mit dem χ^2 -Test nach Pearson zeigte, dass die Nullhypothese der Unabhängigkeit der beiden Verteilungen Diskriminierungsart und Altersgruppen deutlich verworfen werden konnte, $\chi^2(8, N = 140) = 22.61; p = .01$.

Diskriminierungsleistung in Abhängigkeit von der prädiktiven Regelanwendung

Da bereits feststeht, dass die Interaktion von Art des Mischungsresultates und prädiktiver Regelanwendung bei den Realitätseinschätzungen signifikant war (gleiche Intensitätskombination, S. 111; unterschiedliche Intensitätskombination, S. 114), kann davon ausgegangen werden, dass die Art der prädiktiven Regelanwendung auch einen Effekt auf die Diskriminierungsfähigkeit hatte. Da aber die Wechselwirkung von Art des Mischungsresultates, Regelanwendung und Alter vorher nicht signifikant war, waren auch keine Interaktionen von Regel und Alter bei der Diskriminierungsleistung zu erwarten. Die Berechnung der zweifaktoriellen univariaten Varianzanalyse mit der Diskriminierungsleistung zeigte denn auch erwartungsgemäss bei beiden Intensitätskombinationen

signifikante Haupteffekte der beiden Zwischensubjektfaktoren Alter und prädiktive Regelanwendung (ohne Restkategorie)⁵.

Gleiche Intensitätskombination: Entwicklungsverläufe der verschiedenen Regelanwender

Mit je einer univariaten Varianzanalyse mit dem Faktor Alter wurde der Entwicklungsverlauf für jede der drei Kategorien der prädiktiven Regelanwendung überprüft. Der Alterseffekt war bei den Durchschnittsregelanwender höchst signifikant, $F(4,74)=10.36$, $p < .001$. Die Post-Hoc-Tests nach Scheffé errechneten zwischen den Sechsjährigen und allen anderen Altersgruppen signifikante Unterschiede in der Höhe der Diskriminierung. Bei den Additions- und inkonsistenten Regelanwendern war der Faktor Alter nicht signifikant, was der geringen Stichprobengröße zuzuschreiben ist ($n = 27$ bzw. $n = 34$). Die varianzanalytische Überprüfung eines Effektes der Regelanwendung in jeder Altersgruppe ergab bei den Zehnjährigen ($F(2,28) = 3.34$, $p = .052$) und den Zwölfjährigen ($F(2,30) = 2.87$, $p = .074$) je einen Trend. Die Post-Hoc-Tests nach Scheffé wiesen darauf hin, dass diese Effekte durch die Unterschiede zwischen Additions- und Durchschnittsregelanwender zu Stande kamen.

Die Entwicklungsverläufe der Diskriminierungsleistungen der verschiedenen Regelanwender sind in Abbildung 3.3.13 dargestellt.

⁵ Sowohl bei der gleichen als auch bei der unterschiedlichen Intensitätskombination waren die Faktoren Alter ($F(4,136) = 6.88$, $p < .001$ bzw. $F(4,139) = 6.17$, $p < .001$) und konsistente Regelanwendung ($F(2,136) = 4.57$, $p < .05$ bzw. $F(2,139) = 4.24$, $p < .05$) höchst signifikant bzw. signifikant. Bei der gleichen Intensitätskombination gelang die Diskriminierung in der Reihenfolge Durchschnittsbilder > Inkonsistente > Addierer. Die Post-Hoc-Tests nach Scheffé berechneten signifikante Unterschiede zwischen Durchschnitts- und inkonsistenten Regelanwendern (mittlere Differenz = 2.29, $p < .01$) und zwischen Durchschnitts- und Additionsregelanwendern (mittlere Differenz = 3.75, $p < .001$). Bei der unterschiedlichen Intensitätskombination lag die Reihenfolge Durchschnittsbilder > Addierer > Inkonsistente vor. Auch hier zeigten die Post-Hoc-Tests nach Scheffé, dass sich die Durchschnitts- und Additionsregelanwender (mittlere Differenz = 2.7, $p < .01$) und die Durchschnitts- und inkonsistenten Regelanwender (mittlere Differenz = 3.31, $p < .001$) statistisch bedeutsam unterschieden.

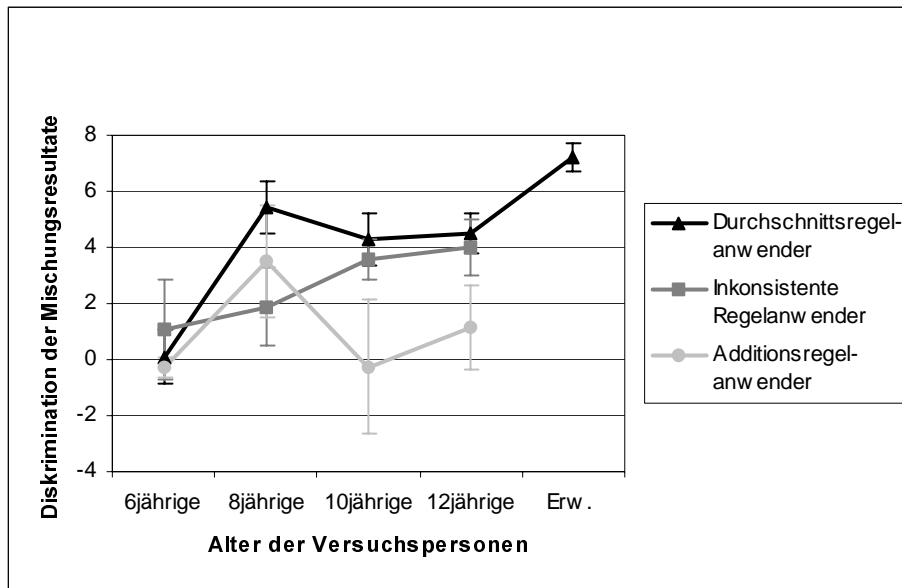


Abbildung 3.3.13. Mittelwerte und Standardfehler der Diskriminationen zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen bei der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Regelanwendung und der Altersgruppe. Die Ausprägungen der Diskriminierungen sind definiert durch die Differenz der Einschätzungen mit der Zauberskala zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen, d. h. positive Diskriminierungen entsprechen höheren Werten mit der Zauberskala bei den physikalisch unkorrekten Mischungen als bei den korrekten Mischungen und somit korrekten Realitätseinschätzungen.

Abbildung 3.3.13 zeigt, dass die Durchschnittsregelanwender – mit Ausnahme der Sechsjährigen – stets besser diskriminierten als die anderen Regelanwender. Allerdings nimmt das Ausmass der Diskriminierung bei den Kindern nicht stetig zu, sondern war bei den Zehn- und Zwölfjährigen im Gegensatz zu den Achtjährigen leicht geringer. Interessant ist auch der Verlauf der Unterscheidung der beiden Mischungsarten der gleichen Intensitätskombination bei den prädiktiv inkonsistenten Regelanwendern: Sie unterschieden die beiden Mischungsarten mit zunehmendem Alter kontinuierlich besser. Bei den Additionsregelanwendern aus Phase 1 hingegen diskriminierten die älteren Kinder durchschnittlich fast kaum mehr.

Gleiche Intensitätskombination: Häufigkeiten

Die Verteilungen der Versuchspersonen auf die drei Kategorien physikalisch korrekte, indifferente und unkorrekte Einschätzung konnten in Bezug zur prädiktiven Regelanwendung gesetzt werden. In Tabelle 3.3.1 sind die Verteilungen in Häufigkeiten und Prozenten dargestellt.

Tabelle 3.3.1

Darstellung der Häufigkeiten und Prozente korrekter, indifferenter und unkorrekter Diskriminierungen der beiden Mischungsarten pro Art der prädiktiven Regelanwendung bei der gleichen Intensitätskombination

Prädiktiver Regeltyp	Diskriminierung					
	korrekt		indifferent		unkorrekt	
Durchschnitt	63	(84.0%)	6	(8.0%)	6	(8.0%)
Inkonsistenz	24	(70.6%)	4	(11.8%)	6	(17.6%)
Addition	11	(40.7%)	7	(25.9%)	9	(33.3%)
Alle Vpn (inkl. Restkategorie)	100	(71.4%)	19	(13.6%)	21	(15.0%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Die Ausgabe der standardisierten Residuen mittels Kreuztabelle wies auf die folgenden höchsten Abweichungen: viele unkorrekte und indifferente Addierer (standardisiertes Residuum = 2.4 bzw. 2.0) sowie wenig korrekte Addierer (standardisiertes Residuum = -1.9). Die Berechnung mit dem χ^2 -Test nach Pearson ergab, dass die beiden Verteilungen deutlich voneinander abhängig waren, $\chi^2(4, N = 136) = 18.69; p = .001$. Allerdings ist dadurch, dass 33% der Zellen eine erwartete Häufigkeit < 5 aufwiesen, der Wert nur bedingt interpretierbar. Deshalb wurden die Verteilungen der korrekten und unkorrekten Durchschnitts- und Additionsregelanwender zusätzlich noch mit dem Vierfelder- χ^2 -Test abgesichert (Fishers exakter Test, $p < .001$).

In Tabelle 3.3.2 sind die Häufigkeiten der drei Diskriminierungskategorien pro Regelanwendungskategorie und Altersgruppe abgebildet.

Tabelle 3.3.2

Häufigkeiten korrekter, indifferenter und unkorrekter Diskriminierungen der beiden Mischungsarten pro Altersgruppe und Art der prädiktiven Regelanwendung bei der gleichen Intensitätskombination

	Prädiktive Regelanwendung								
	Durchschnitt			Inkonsistent			Addition		
	korrekt	Indiff.	unkorr.	korrekt	Indiff.	unkorr.	korrekt	Indiff.	unkorr.
Sechsjährige (n=30)	4	3	4	5	2	3	1	4	2
Achtjährige (n=30)	13	1	1	5	1	3	3	1	2

Zehnjährige (n=30)	9	1		11	1		3		3
Zwölfjährige (n=29)	17	1	1	3			4	2	2
Erwachsene (n=20)	20								

Unterschiedliche Intensitätskombination: Entwicklungsverläufe der verschiedenen Regelanwender

Mit je einer univariaten Varianzanalyse mit dem Faktor Alter wurde der Entwicklungsverlauf für jede der drei Kategorien der prädiktiven Regelanwendung überprüft. Der Alterseffekt war bei den Durchschnittsregelanwendern höchst signifikant, $F(4,87) = 6.66$, $p < .001$. Die Überprüfung mit dem Post-Hoc-Test nach Scheffé lokalisierte signifikante Altersunterschiede zwischen den Sechsjährigen und den Zwölfjährigen bzw. den Erwachsenen. Bei den Additions- und inkonsistenten Regelanwendern war der Faktor Alter nicht signifikant, was der geringen Stichprobengrösse zuzuschreiben ist ($n = 24$ bzw. $n = 27$).

Die varianzanalytische Überprüfung eines Effektes der Regelanwendung in jeder Altersgruppe ergab bei den Zehnjährigen ($F(2,30) = 3.6$, $p < .05$) und den Zwölfjährigen ($F(2,30) = 4.9$, $p < .05$) je einen signifikanten Haupteffekt. Der Post-Hoc-Tests nach Scheffé wies bei den Zwölfjährigen auf einen sehr hohen Unterschied zwischen Additions- und Durchschnittsregelanwender hin (mittlere Differenz = 6.98, $p < .05$).

Die Entwicklungsverläufe der Diskriminierungsleistungen bei der unterschiedlichen Intensitätskombination sind in Abbildung 3.3.14 für die verschiedenen Regelanwender dargestellt.

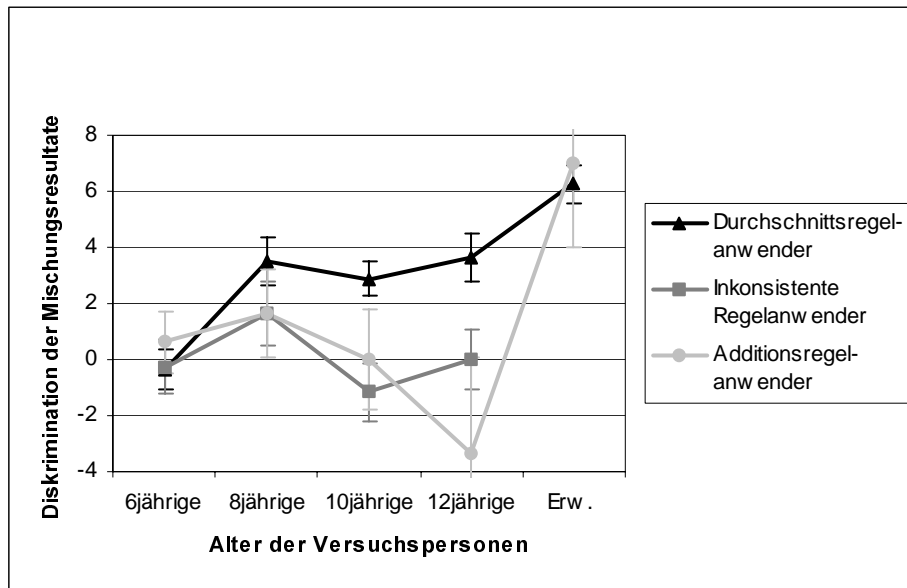


Abbildung 3.3.14. Mittelwerte und Standardfehler der Diskriminationen zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen bei der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Regelanwendung und der Altersgruppe. Die Ausprägungen der Diskriminierungen sind definiert durch die Differenz der Einschätzungen mit der Zauberskala zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen, d. h. positive Diskriminierungen entsprechen höheren Werten mit der Zauberskala bei den physikalisch unkorrekten Mischungen als bei den korrekten Mischungen und somit korrekten Realitätseinschätzungen.

Vergleicht man die obige Darstellung der mittleren Diskriminierungen der verschiedenen Regelanwender bei der unterschiedlichen Intensitätskombination mit derjenigen bei der gleichen Intensitätskombination (Abbildung 3.3.13), so fällt auf, dass die Durchschnittsregelanwender bei beiden Bedingungen sehr ähnlich einschätzten. Auch bei der unterschiedlichen Intensitätskombination war die Diskriminierungsfähigkeit bei den Achtjährigen schon recht hoch und verbesserte sich bei den älteren Kindern nicht mehr. Erst bei den Erwachsenen war eine deutlich bessere Diskriminierungsfähigkeit zu beobachten. Der Verlauf der inkonsistenten Regelanwender wies hier aber nicht auf eine zunehmende Realitätsanpassung, sondern zeigte vielmehr, dass die beiden Mischungen kaum mehr unterschieden wurden. Bei den Addierern verdeutlichte sich bei der unterschiedlichen Intensitätskombination die negative Diskriminierung mit zunehmendem Alter. Die zwölfjährigen Addierer unterscheiden die beiden Mischungsarten so deutlich wie die gleichaltrigen Durchschnittsregelanwender, nur in umgekehrter Richtung.

Unterschiedliche Intensitätskombination: Häufigkeiten

Die Werte der Diskriminierungen bei der unterschiedlichen Intensitätskombination wurden wie folgt kategorisiert: korrekte (positive Werte), unkorrekte (negative Werte)

und indifferente (Wert null) Einschätzung. Die Verteilungen der Einschätzungen aus Phase 1 sind in Tabelle 3.3.3 in Abhängigkeit von der Art der Diskrimination dargestellt.

Tabelle 3.3.3

Darstellung der Häufigkeiten und Prozente korrekter, indifferenter und unkorrekter Diskriminierungen der beiden Mischungsarten pro Art der prädiktiven Regelanwendung bei der unterschiedlichen Intensitätskombination

Prädiktiver Regeltyp	Diskriminierung					
	korrekt		indifferent		unkorrekt	
Durchschnitt	64	(72.7%)	16	(18.2%)	8	(9.1%)
Inkonsistenz	10	(37.0%)	6	(22.2%)	11	(40.7%)
Addition	10	(41.7%)	4	(16.7%)	10	(41.7%)
Alle Vpn (inkl. Restkategorie)	84	(60.0%)	27	(19.3%)	29	(20.7%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Der χ^2 -Test nach Pearson zeigte, dass die beiden Verteilungen höchst signifikant voneinander abhängig waren, $\chi^2(4, N = 139) = 22.23$; $p < .001$. Die Berechnung der standardisierten Residuen wies bei den unkorrekten Durchschnittsbildern (wenige; standardisiertes Residuum = -2.4), inkonsistenten Regelanwendern und Addierern (viele; standardisiertes Residuum = 2.3 bzw. 2.2) auf die höchsten Abweichungen.

In Tabelle 3.3.4 sind die Häufigkeiten der drei Diskriminierungskategorien pro Regelanwendungskategorie und Altersgruppe abgebildet.

Tabelle 3.3.4

Häufigkeiten korrekter, indifferenter und unkorrekter Diskriminierungen der beiden Mischungsarten pro Altersgruppe und Art der prädiktiven Regelanwendung bei der unterschiedlichen Intensitätskombination

Altersgruppe	Prädiktive Regelanwendung								
	Durchschnitt			Inkonsistent			Addition		
	korrekt	Indiff.	unkorr.	korrekt	Indiff.	unkorr.	korrekt	Indiff.	unkorr.
Sechsjährige (n=30)	2	3	4	4	3	5	3	2	3
Achtjährige (n=30)	11	3	1	4	2	2	3	1	3

Zehnjährige (n=30)	18	4	1	1		2	1	1	2
Zwölfjährige (n=29)	16	5	2	1	1	2	1		2
Erwachsene (n=20)	17	1					2		

Subjektive Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsergebnisse

Die Erfassung der subjektiven Sicherheit des prädiktiven Mischungsurteils nach der Beobachtung des Mischungsvorganges erwies sich bei den Sechsjährigen als problematisch, da der Eindruck entstand, als würden sie die Skala nicht verstehen. Im Verlauf der Versuchsdurchführung wurde deshalb auf die Erfassung bei den Sechsjährigen verzichtet. In diesem Kapitel wurden die Daten der wenigen Sechsjährigen deshalb nicht mit einbezogen. Zudem fehlte bei einer Zwölfjährigen die Erfassung bei den Versuchsdurchgängen gf und uf. Somit bestand die gesamte Stichprobe bei den physikalisch unkorrekten Mischungsergebnissen aus 109 und bei den physikalisch korrekten Mischungsergebnissen aus 110 Versuchspersonen.

Ergebnisse der Varianzanalyse

Die Varianzanalyse mit den Messwiederholungsfaktoren Farbintensitätskombination und Art des Mischungsergebnisses sowie den Zwischensubjektfaktoren Alter und Geschlecht ergab die beiden Haupteffekte der Faktoren Art des Mischungsergebnisses ($F(1, 101) = 5.07, p < .05$) und Alter ($F(3, 101) = 8.65, p < .001$). Die Intensitätskombination wirkte sich hier also nicht direkt auf die abhängige Variable aus. Erst bei der Dreifachwechselwirkung der Faktoren Mischungsergebnis, Alter und Intensitätskombination war ein Effekt der verschiedenen Ausgangsintensitäten festzustellen ($F(3, 101) = 2.56, p = .059$). Die Bedeutung dieser Wechselwirkung kann anhand der Abbildung 3.3.15, welche die subjektive Sicherheit der prädiktiven Mischungsregel nach Rückmeldung in Abhängigkeit von Mischungsergebnis, Farbintensitätskombination und Alter darstellt, abgelesen werden: Die Zwölfjährigen zeigten zwischen den beiden Mischungsergebnissen der unterschiedlichen Intensitätskombination und die Achtjährigen bei der gleichen Intensitätskombination einen sehr deutlichen Unterschied, während sich in den anderen

Altersgruppen keine hohen Unterschiede zwischen den Mischungsergebnissen zeigten. Die Varianzanalyse ergab auch hier, dass das Geschlecht der Versuchspersonen keinen Einfluss auf die Regelsicherheit nach der Beobachtung der Mischungsvorgänge hatte. Auch die Überprüfung eines Reihenfolgeeffektes zeigte keine statistische Bedeutung der Reihenfolgen der Versuchsdurchgänge.

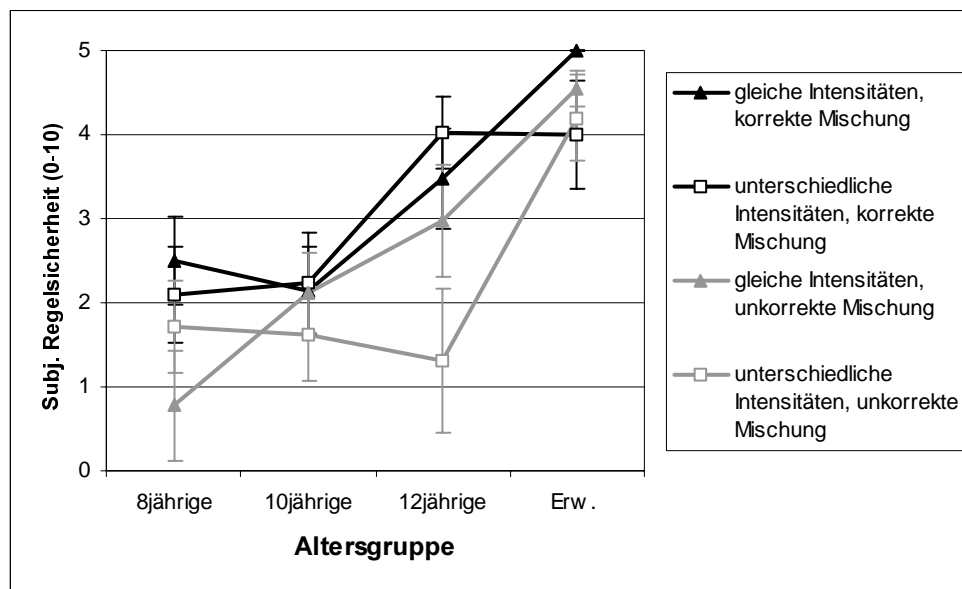


Abbildung 3.3.15. Mittelwerte und Standardfehler der subjektiven Sicherheit der prädiktiven Mischungsregel nach der Rückmeldung durch die physikalisch korrekten und unkorrekten Mischungsergebnisse der vier Versuchsdurchgänge. Werte > 0 bedeuten eine Beibehaltung der ursprünglichen Regel aus Phase 1. Werte von null besagen, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Der Abbildung 3.3.15 lässt sich entnehmen, dass sich die Mischungsergebnisse auf die subjektive Regelsicherheit i.S. tieferer Werte bei den physikalisch unkorrekten Mischungsergebnissen auswirkten (graue Linien). Die Berechnungen von Varianzanalysen pro Altersgruppe zeigten, dass sich das Mischungsergebnis aber nur bei den Zwölfjährigen signifikant auswirkte ($F(1, 28) = 4.74, p < .05$). Der Post-Hoc-Test nach Scheffé zur Überprüfung des Alterseffektes auf die mittlere Sicherheit ergab, dass sich die Altersgruppen der Acht- und Zehnjährigen je von den Erwachsenen signifikant ($p < .001$ bzw. $p = .001$) bzw. bei den Zwölfjährigen mit einem Trend ($p = .064$) unterschieden. Die subjektive Sicherheit stieg demnach bei allen vier Versuchsdurchgängen zusammen mit zunehmendem Alter stetig an.

Subjektive Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsergebnisse in Abhängigkeit von der Regel

Gleiche Intensitätskombination: Mittelwerte

Die Kovarianzanalyse mit dem Messwiederholungsfaktor Mischungsergebnis, den Zwischensubjektfaktoren Alter und konsistente Regelanwendung sowie der Kovariate mittlere Regelsicherheit vor dem Mischungsvorgang zeigte, dass die generelle Sicherheit des prädiktiven Mischungsurteils keinen Einfluss auf die Sicherheit nach der Rückmeldung durch die verschiedenen Mischungsergebnisse hatte.

Die Varianzanalyse mit dem Messwiederholungsfaktor Art des Mischungsergebnisses und den Zwischensubjektfaktoren Alter und konsistente Regelanwendung (ohne die Kategorie der Subtraktionsregel) erbrachte einen höchst signifikanten Haupteffekt des Faktors Alter ($F(3,98) = 7.51, p < .001$). Weiter trat eine höchst signifikante Zweifachwechselwirkung zwischen den Faktoren Art des Mischungsergebnisses und konsistente Regelanwendung auf ($F(2,98) = 15.87, p < .001$; siehe Abbildung 3.3.16).

Aufgrund der Interaktion zwischen der Art des Mischungsergebnisses und der prädiktiven Regelanwendung wurden weitere Varianzanalysen pro Regelanwendungskategorie mit dem Messwiederholungsfaktor Art des Mischungsergebnisses gerechnet. Diese hatten ergeben, dass sich die Sicherheitseinschätzungen bei den verschiedenen Mischungsergebnissen sowohl bei den Durchschnitts- als auch Additionsregelanwendern höchst signifikant unterschieden ($F(1,62) = 15.63, p < .001$ bzw. $F(1,19) = 14.48, p = .001$).

In Abbildung 3.3.16 ist gut zu erkennen, dass die Durchschnittsbilder und die inkonsistenten Regelanwender nach der Rückmeldung durch das physikalisch korrekte Mischungsergebnis sicherer waren als nach der Rückmeldung durch das unkorrekte Resultat. Die Addierer hingegen waren nach Beobachtung des physikalisch unkorrekten Mischungsergebnisses sicherer als nach Beobachtung des korrekten Resultates, welche sogar zu deutlicher Unsicherheit führte (Wert gegen null).

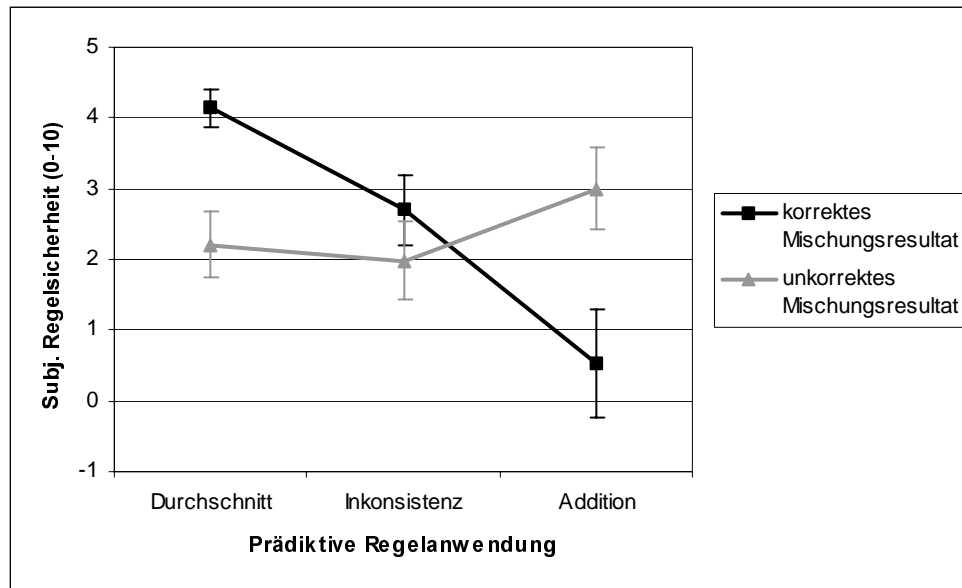


Abbildung 3.3.16. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) in Abhängigkeit von der konsistenten prädiktiven Regelanwendung und der Art des Mischungsresultates bei der gleichen Intensitätskombination. Werte > 0 bedeuten eine Beibehaltung der ursprünglichen Regel aus Phase 1. Werte von null besagten, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Entwicklungsverläufe pro Regelanwendertyp und Mischungsart

Mit je einer univariaten Varianzanalyse mit dem Faktor Alter wurde der Entwicklungsverlauf für jede der drei Kategorien der prädiktiven Regelanwendung überprüft. Bei der physikalisch korrekten Mischungsart (Abbildung 3.3.17) unterschieden sich nur die Durchschnittsregelanwender signifikant, $F(3, 64) = 3.1, p < .05$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich die Achtjährigen und die Erwachsenen bei der Einschätzung mit einem Trend unterschieden ($p = .058$).

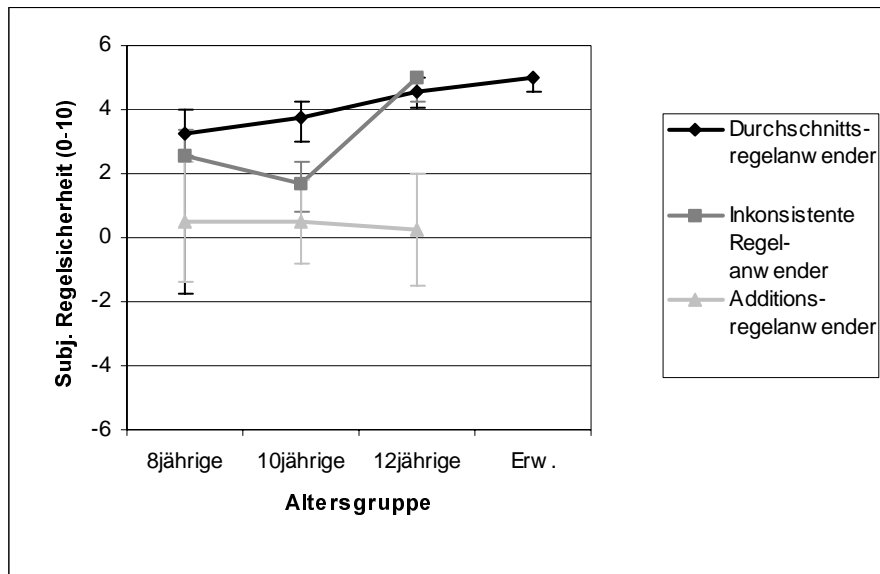


Abbildung 3.3.17. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) nach Beobachtung der physikalisch korrekten Mischungsart der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe. Werte > 0 bedeuten eine Beibehaltung und Werte < 0 eine Verwerfung der ursprünglichen Regel aus Phase 1. Werte von null besagen, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Die Mittelwerte der subjektiven Sicherheitseinschätzungen nach der Beobachtung der Mischungsvorgänge der prädiktiv inkonsistenten Regelanwender sind in Abbildung 3.3.17 (physikalisch korrekte Mischungsart) und Abbildung 3.3.18 (physikalisch unkorrekte Mischungsart) zwar der Vollständigkeit halber mit abgebildet, sie lassen sich allerdings inhaltlich nicht interpretieren.

Der Alterseffekt beim physikalisch unkorrekten Mischungsergebnis (Abbildung 3.3.18) war bei den Durchschnittsregelanwender höchst signifikant, $F(3, 63) = 7.79, p < .001$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich die Achtjährigen und die Erwachsenen bei der Einschätzung signifikant unterschieden ($p < .001$). Auch die Additionsregelanwender unterschieden sich im Entwicklungsverlauf signifikant, $F(2, 20) = 4.88, p < .05$. Der Post-Hoc-Test nach Scheffé wies darauf hin, dass sich die Acht- und Zwölfjährigen signifikant unterschieden ($p < .05$).

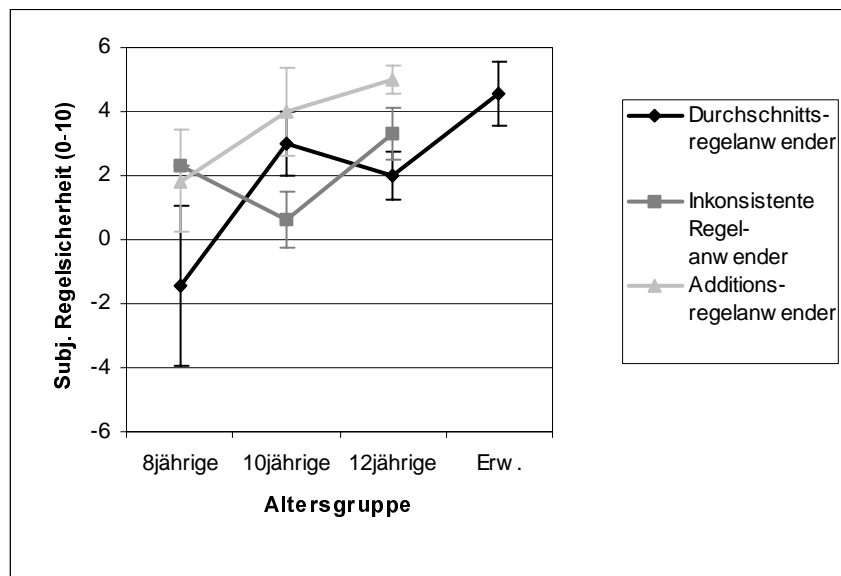


Abbildung 3.3.18. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) nach Beobachtung der physikalisch unkorrekten Mischungsart der gleichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe. Werte > 0 bedeuten eine Beibehaltung und Werte < 0 eine Verwerfung der ursprünglichen Regel aus Phase 1. Werte von null besagten, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Gleiche Intensitätskombination: Häufigkeiten

Die Urteile der subjektiven Sicherheit nach der Rückmeldung durch die Mischungsresultate wurden in drei Kategorien eingeteilt: stabile Regelanwender (positive Werte), verunsicherte Regelanwender (Wert null) und instabile bzw. revidierende Regelanwender (negative Werte). Die drei Umgangsarten mit der prädiktiven Regel nach Beobachtung der Mischungsresultate wurden mit der prädiktiven Regelanwendung selbst in Beziehung gesetzt.

Zunächst werden die kategorisierten Regelsicherheiten nach physikalisch korrektem Mischungsresultat dargestellt (Tabelle unten).

Tabelle 3.3.5

Darstellung der Häufigkeiten und Prozente bei physikalisch korrektem Mischungsresultat bei der gleichen Intensitätskombination

Prädiktiver Regeltyp	Umgang mit der prädiktiven Regel nach Mischungsvorgang		
	stabil	verunsichert	instabil
Durchschnitt	61 (95.3%)	1 (1.6%)	2 (3.1%)

Addition	10	(50%)	4	(20%)	6	(30%)
----------	----	-------	---	-------	---	-------

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Der χ^2 -Test nach Pearson mit den 6 Zellen ergab zwar einen höchst signifikanten χ^2 -Wert, der aber dadurch, dass 50% der Zellen eine erwartete Häufigkeit < 5 hatten, nicht interpretierbar war. Deshalb wurden die Verteilungen der stabilen und instabilen Durchschnitts- und Additionsregelanwender mit dem Vierfelder- χ^2 -Test noch abgesichert (Fishers exakter Test, $p < .001$). Die Kreuztabelle wies die höchsten standardisierten Residuen bei den vielen instabilen und verunsicherten Addierer (standardisiertes Residuum = 3 bzw. 2.6) sowie den wenigen instabilen Durchschnittsbilder (standardisiertes Residuum = -1.7) auf.

Die Überprüfung der Verteilungen bei den kategorisierten Regelsicherheiten nach physikalisch unkorrektem Mischungsresultat und der prädiktiven Regelanwendung mit dem χ^2 -Test nach Pearson ergab kein signifikantes Resultat. Die höchsten Abweichungen waren durch die wenigen instabilen Addierer gegeben (standardisiertes Residuum = -1.2).

Tabelle 3.3.6

Darstellung der Häufigkeiten und Prozente bei physikalisch unkorrektem Mischungsresultat bei der gleichen Intensitätskombination

Prädiktiver Regeltyp	Umgang mit der prädiktiven Regel nach Mischungsvorgang					
	stabil		verunsichert		instabil	
Durchschnitt	46	(73%)	5	(7.9%)	12	(19%)
Addition	17	(85%)	2	(10%)	1	(5%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt. Bei den Durchschnittsbildern fehlt ein Wert. Deshalb ergibt die Summe der Prozente in der Zeile der Durchschnittsbilder nicht 100%.

Die Tabelle 3.3.5 zeigt, dass insgesamt 61 Personen (95 Prozent) der Durchschnittsregelanwender beim physikalisch korrekten Mischungsresultat bei ihrem Urteil blieben und nur 2 (3 Prozent) ihre Meinung änderten. Beim physikalisch unkorrekten Mischungsresultat (Tabelle 3.3.6) blieben immerhin noch 46 Personen (73 Prozent) bei ihrem Urteil und 12 Personen (19 Prozent) revidierten es. 10 Personen (50 Prozent) der Addierer blieben beim physikalisch korrekten Mischungsresultat bei ihrem Urteil, 6 (30 Prozent) änderten beim physikalisch korrekten Mischungsresultat ihre Meinung. Beim

physikalisch unkorrekten Mischungsresultat blieben 17 Personen (85 Prozent) bei ihrem Urteil und nur 1 Person (5 Prozent) revidierten es.

Konsistente Einschätzungen bei der subjektiven Sicherheit

Die beiden Einschätzungen des Umgangs mit der prädiktiven Regel anhand der Sicherheitsskala wurden bezüglich ihrer Konsistenz bei beiden Mischungsresultaten korrekten, inkonsistenten und unkorrekten Einschätzungen zugewiesen. Eine korrekte Einschätzung lag vor, wenn die Durchschnittsregelanwender ihre Regel zweimal bestätigten, wenn die inkonsistenten Regelanwender ihre beiden Regeln konsistent richtig einschätzten (meistens die Additionsregel verwarfen und die Durchschnittsregel bestätigten) und wenn die Additionsregelanwender ihre Regel zweimal verwarfen. Die Zuweisung der inkonsistenten und konsistent unkorrekten Einschätzungen erfolgte analog. In Abbildung 3.3.19 sind die prozentualen Anteile der drei Kategorien pro Altersgruppe aufgetragen.

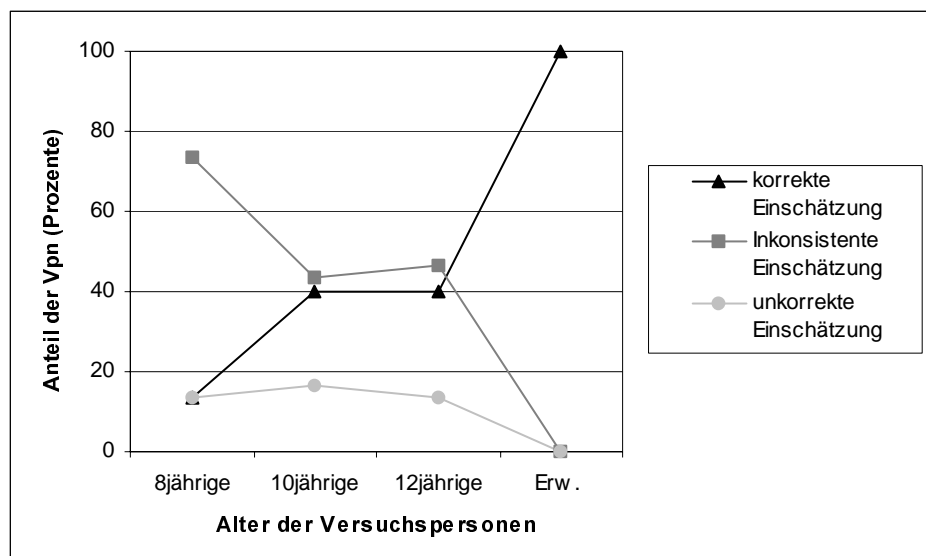


Abbildung 3.3.19. Konsistent korrekte und unkorrekte sowie inkonsistente Einschätzungen im Hinblick auf die prädiktive Regelanwendung anhand der subjektiven Sicherheit bei der gleichen Intensitätskombination in Prozenten pro Altersgruppe.

Die Überprüfung der Verteilungen der beiden Variablen auf Unabhängigkeit mit dem χ^2 -Test nach Pearson zeigte einen höchst signifikanten Wert, $\chi^2(6, N = 110) = 38.66$; $p < .001$. Da aber vier Zellen (33.3%) eine erwartete Häufigkeit kleiner 5 aufwiesen, wurde ein erneuter Test ohne die dritte Kategorie der konsistent unkorrekten Einschätzungen durchgeführt, $\chi^2(3, N = 103) = 32.65$; $p < .001$.

In Tabelle 3.3.7 sind die Häufigkeiten und prozentualen Anteile der verschiedenen Einschätzungen schematisch dargestellt.

Tabelle 3.3.7

Darstellung der Häufigkeiten und Prozente der konsistenten Einschätzungen bei der physikalisch unmöglichen und möglichen Mischungsart mit der Sicherheitsskala nach Beobachtung der Mischungsvorgänge bei der gleichen Intensitätskombination

Prädiktiver Regeltyp	Konsistenz der Urteile beim bewussten Umgang mit der Regel					
	korrekt		inkonsistent		unkorrekt	
Durchschnitt	45	(70.3%)	18	(28.1%)	1	(1.6%)
Inkonsistenz	3	(12.5%)	19	(79.2%)	2	(8.3%)
Addition	0	(0%)	9	(45%)	11	(55%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Um eine Aussage über den Entwicklungsverlauf des konsistenten Umgangs mit der prädiktiven Regel nach Beobachtung der Mischungsvorgänge machen zu können, wurden die Häufigkeiten konsistenter korrekter und unkorrekter sowie inkonsistenter Einschätzungen pro Altersgruppe in Tabelle 3.3.8 aufgeführt.

Tabelle 3.3.8

Häufigkeiten konsistenter korrekter und unkorrekter sowie indifferenter Einschätzungen mit der Sicherheitsskala nach Beobachtung der Mischungsvorgänge pro Altersgruppe und Art der prädiktiven Regelanwendung bei der gleichen Intensitätskombination

Altersgruppe	Prädiktive Regelanwendung								
	Durchschnitt			Inkonsistent			Addition		
	korrekt	inkons.	unkorr.	korrekt	inkons	unkorr.	korrekt	inkons.	unkorr.
Achtjährige (n=30)	4	11		1	8		3		3
Zehnjährige (n=28)	9	1		1	8	3	3		3
Zwölfjährige (n=30)	12	6	1		3		3		5
Erwachsene (n=20)	20								

Die Verteilungen der Häufigkeiten bei der Durchschnittsregel auf die verschiedenen Altersgruppen zeigen deutlich, wie die korrekte Einschätzung mit zunehmendem Alter häufiger wurde. Insgesamt liess sich nur eine Person konsistent umstimmen. Bei den

inkonsistenten Einschätzungen der Durchschnittsregelanwender aus Phase 1 widerspiegeln sich die Verunsicherungen durch das physikalisch unmögliche Mischungsresultat (vgl. Abbildung 3.3.18): Die Acht- und Zwölfjährigen liessen sich eher verunsichern, was zu inkonsistenten Urteilen führte.

Bei den Verteilungen der Häufigkeiten der prädiktiven Addierer fällt zunächst auf, dass kein einziges Kind die falsche Additionsregel konsistent verwerfen konnte. Die Hälfte der Addierer urteilte inkonsistent und die andere Hälfte konsistent unkorrekt, d. h. sie blieben ihrer Regel treu. Bei den Zwölfjährigen war die Additionsregel sogar mehrheitlich stabil.

Unterschiedliche Intensitätskombination: Mittelwerte

Die Kovarianzanalyse mit dem Messwiederholungsfaktor Art des Mischungsresultates, den Zwischensubjektfaktoren Alter und konsistente Regelanwendung sowie der Kovariate mittlere Regelsicherheit vor dem Feedback zeigte, dass die generelle Sicherheit des prädiktiven Mischungsurteils keinen Einfluss auf die Sicherheit nach dem Feedback hatte.

Die Varianzanalyse mit dem Messwiederholungsfaktor Art des Mischungsresultates und den Zwischensubjektfaktoren Alter und konsistente Regelanwendung (ohne die Kategorie der Zentrierungsregel) erbrachte signifikante Haupteffekte der Faktoren Alter ($F(3,98) = 3.06, p < .05$) und konsistente Regelanwendung ($F(2,98) = 14.52, p < .001$). Weiter traten höchst signifikante Zweifachwechselwirkungen zwischen den Faktoren Art des Mischungsresultates und konsistente Regelanwendung ($F(2,98) = 12.9, p < .001$) sowie zwischen Alter und konsistenter Regelanwendung ($F(5,98) = 5.54, p < .001$) auf.

Bedingt durch die Zweifachwechselwirkungen zwischen den Faktoren Art des Mischungsresultates und konsistente Regelanwendung wurden einzelne Varianzanalysen mit der Art des Mischungsresultates als Messwiederholungsfaktor für jede Regelanwendungskategorie gerechnet. Es zeigte sich, dass sich die Sicherheitseinschätzungen der verschiedenen Mischungsresultate sowohl bei den Durchschnitts- als auch Additionsregelanwendern höchst signifikant unterschieden ($F(1,77) = 20.19, p < .001$ bzw. $F(1,15) = 15.77, p = .001$). Abbildung 3.3.20 veranschaulicht diesen Sachverhalt.

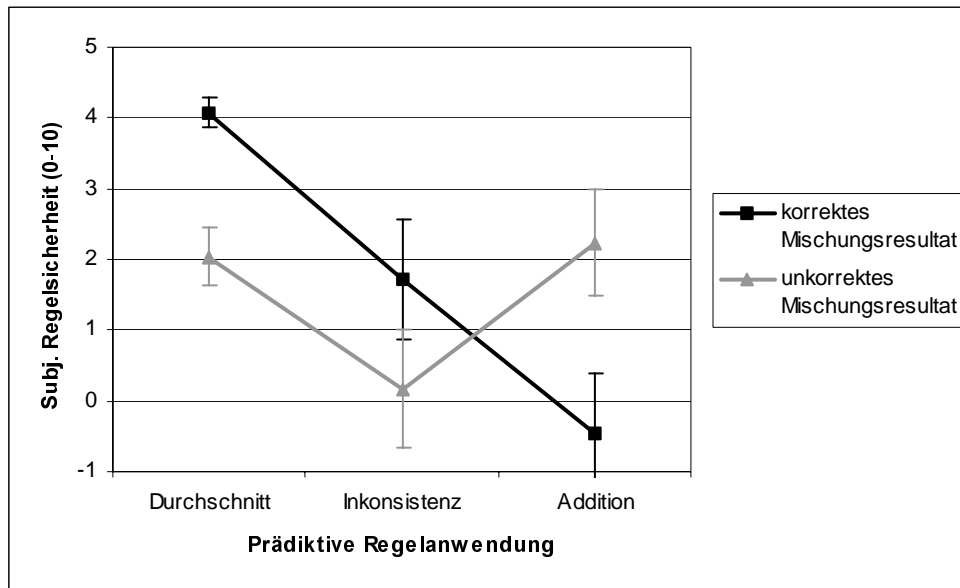


Abbildung 3.3.20. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) in Abhängigkeit von der konsistenten prädiktiven Regelanwendung und der Art des Mischungsresultates bei der unterschiedlichen Intensitätskombination. Werte > 0 bedeuten eine Beibehaltung und Werte < 0 eine Verwerfung der ursprünglichen Regel aus Phase 1. Werte von null besagen, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Bedingt durch die Zweifachwechselwirkungen zwischen den Faktoren Alter und konsistente Regelanwendung wurden weitere Varianzanalysen, getrennt für jede Altersgruppe, gerechnet. Es resultierte ein höchst signifikanter Haupteffekt der Regelanwendung bei den Zehnjährigen ($F(2,27) = 10.45, p < .001$) und den Erwachsenen ($F(1,18) = 960.4, p < .001$). Der Post-Hoc-Test nach Scheffé bei den Zehnjährigen wies sowohl zwischen Durchschnitts- und Additionsregelanwendern ($p < .001$) als auch zwischen Durchschnitts- und inkonsistenten Regelanwendern ($p < .05$) signifikante Unterschiede auf. Bei den Erwachsenen bezog sich die Angabe des Haupteffektes bereits auf den Unterschied zwischen den Durchschnitts- und Additionsregelanwender, da keine erwachsene Person in Phase 1 inkonsistent geurteilt hatte. Bei beiden Altersgruppen waren die Mittelwerte der Durchschnittsregelanwender stets am höchsten.

Entwicklungsverläufe pro Regelanwendertyp und Mischungsart

Mit je einer univariate Varianzanalyse mit dem Faktor Alter wurde der Entwicklungsverlauf für jede der drei Kategorien der prädiktiven Regelanwendung überprüft. Bei der physikalisch korrekten Mischungsart (Abbildung 3.3.21) war der Alterseffekt bei den Durchschnittsregelanwendern höchst signifikant, $F(3, 79) = 6.4, p = .001$. Der Post-

Hoc-Test nach Scheffé zeigte, dass sich die Acht- und Zehnjährigen je von den Erwachsenen bei der Einschätzung signifikant unterschieden ($p < .01$). Die Additionsregelanwender wiesen trotz kleiner Stichprobengrösse einen Trend beim Effekt des Alters auf, $F(3, 16) = 3.22, p = .061$.

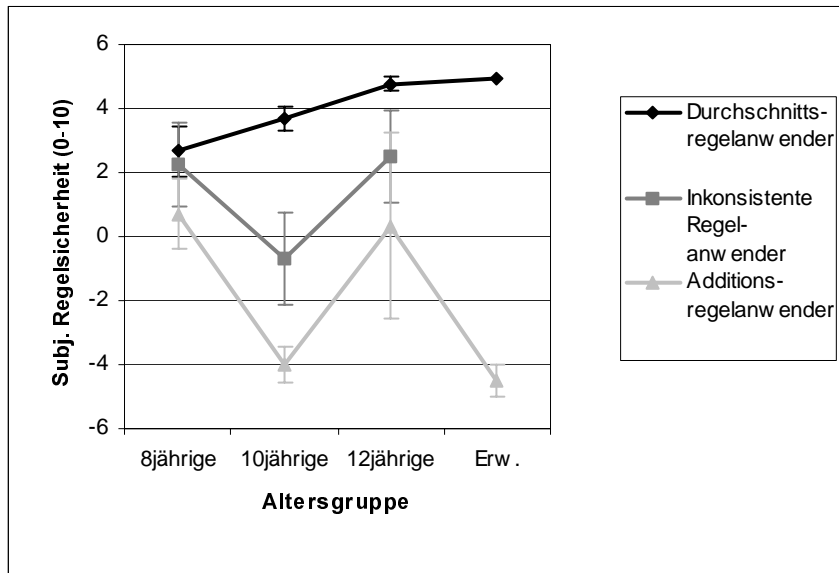


Abbildung 3.3.21. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) nach Beobachtung der physikalisch korrekten Mischungsart der unterschiedlichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe. Werte > 0 bedeuten eine Beibehaltung und Werte < 0 eine Verwerfung der ursprünglichen Regel aus Phase 1. Werte von null besagten, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Bei der physikalisch unkorrekten Mischungsart (Abbildung 3.3.22) unterschieden sich die Durchschnittsregelanwender bezüglich Alter sehr signifikant, $F(3, 78) = 5.37, p < .01$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich alle Altersgruppen der Kinder je von den Erwachsenen bei der Einschätzung in statistisch bedeutsamem Ausmass unterschieden (Achtjährige und Erwachsene, $p < .05$; Zehnjährige und Erwachsene, $p = .073$; Zwölfjährige und Erwachsene, $p < .01$). Auch die Additionsregelanwender unterschieden sich – trotz der geringen Stichprobengrösse – signifikant, $F(3, 16) = 4.77, p < .05$. Der Post-Hoc-Test nach Scheffé zeigte, dass sich die Zwölfjährigen und die Erwachsenen signifikant unterschieden ($p < .05$).

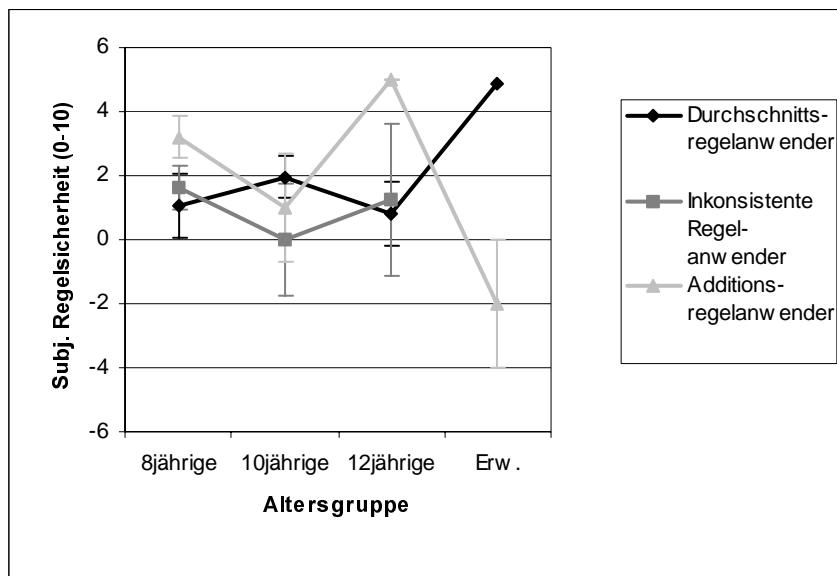


Abbildung 3.3.22. Mittelwerte und Standardfehler der Einschätzungen auf der Skala der subjektiven Sicherheit der prädiktiven Regelanwendung (-5 bis 5) nach Beobachtung der physikalisch unkorrekten Mischungsart der unterschiedlichen Intensitätskombination in Abhängigkeit von der Art der prädiktiven Mischungsregelanwendung und von der Altersgruppe. Werte > 0 bedeuten eine Beibehaltung und Werte < 0 eine Verwerfung der ursprünglichen Regel aus Phase 1. Werte von null besagten, dass die ursprüngliche Regel weder beibehalten noch verworfen werden konnte.

Unterschiedliche Intensitätskombination: Häufigkeiten

Die Verteilungen der kategorisierten subjektiven Sicherheitseinschätzungen nach der Rückmeldung durch die Mischungsresultate in stabile, verunsicherte und instabile bzw. revidierende Regelanwender konnte zusammen mit den Verteilungen der prädiktiven Regelanwendung anhand einer 2x3-Kreuztabelle dargestellt werden (Tabelle 3.3.9)

Tabelle 3.3.9

Darstellung der Häufigkeiten und Prozente bei physikalisch korrektem Mischungsresultat bei der unterschiedlichen Intensitätskombination

Prädiktiver Regeltyp	Umgang mit der prädiktiven Regel nach Mischungsvorgang					
	stabil		verunsichert		instabil	
Durchschnitt	73	(92%)	4	(5%)	2	(3%)
Addition	5	(31%)	2	(13%)	9	(56%)

Anmerkung: In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Der χ^2 -Test nach Pearson zum Vergleich der beiden Verteilungen wies einen höchst signifikanten χ^2 -Wert auf. Da aber die Voraussetzungen des Tests dadurch, dass 50% der Zellen eine erwartete Häufigkeit < 5 hatten, nicht erfüllt waren, darf das Resultat

nicht interpretiert werden. Deshalb wurden die Verteilungen der stabilen und instabilen Durchschnitts- und Additionsregelanwender mit dem Vierfelder- χ^2 -Test überprüft (Fishers exakter Test, $p < .001$). Die 2x3-Kreuztabelle wies folgende bedeutsamen Abweichungen auf: Viele instabile und wenige stabile Addierer (standardisiertes Residuum = 5.3 bzw. -2.2) sowie wenige instabile Durchschnittsbilder (standardisiertes Residuum = -2.4).

Wie schon bei der gleichen Intensitätskombination ergab die Überprüfung der Verteilungen der beiden Variablen beim physikalisch unkorrekten Mischungsresultat (Häufigkeiten und Prozente s. Tabelle 3.3.10) mit dem χ^2 -Test nach Pearson auch bei der unterschiedlichen Intensitätskombination kein signifikantes Resultat.

Tabelle 3.3.10

Darstellung der Häufigkeiten und Prozente bei physikalisch unkorrektem Mischungsresultat bei der unterschiedlichen Intensitätskombination

Prädiktiver Regeltyp	Umgang mit der prädiktiven Regel nach Mischungsvorgang					
	stabil		verunsichert		instabil	
Durchschnitt	53	(68%)	8	(10%)	17	(22%)
Addition	11	(69%)	3	(19%)	2	(13%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt. Bei den Durchschnittsbildern fehlt ein Wert. Deshalb ergibt die Summe der Prozente in der Zeile der Durchschnittsbilder nicht 100%.

Der Tabelle 3.3.9 ist zu entnehmen dass insgesamt 73 Personen (92 Prozent) der Durchschnittsregelanwender bei der physikalisch korrekten Mischung bei ihrem Urteil blieben und nur gerade 2 Personen (3 Prozent) ihre Meinung änderten. Bei der unkorrekten Mischung blieben immerhin noch 53 Personen (68 Prozent) bei ihrem Urteil und 17 (22 Prozent) revidierten es (Tabelle 3.3.10). 5 Personen (31 Prozent) der Addierer blieben bei der Rückmeldung durch das physikalisch korrekte Mischungsresultat bei ihrem Urteil, der grössere Teil, nämlich 9 (56 Prozent) änderten aber ihre Meinung. Bei der physikalisch unkorrekten Mischung hielten 11 Personen (69 Prozent) an ihrer unkorrekten Regel fest und nur 2 Personen (13 Prozent) revidierten sie.

Konsistente Einschätzungen bei der subjektiven Sicherheit

Die beiden Einschätzungen zum bewussten Umgang mit der prädiktiven Regel anhand der Sicherheitsskala wurden bezüglich ihrer Konsistenz bei beiden Mischungsresultaten

korrekten, inkonsistenten und unkorrekten Einschätzungen zugewiesen. Eine korrekte Einschätzung lag vor, wenn die Durchschnittsregelanwender ihre Regel zweimal bestätigten, wenn die inkonsistenten Regelanwender ihre beiden Regeln konsistent richtig einschätzten (meistens die Additionsregel verwarfen und die Durchschnittsregel bestätigten) und wenn die Additionsregelanwender ihre Regel zweimal verwarfen. Die Zuweisung der inkonsistenten und konsistent unkorrekten Einschätzungen erfolgte analog. In Abbildung 3.3.23 sind die prozentualen Anteile der drei Kategorien pro Altersgruppe abgetragen.

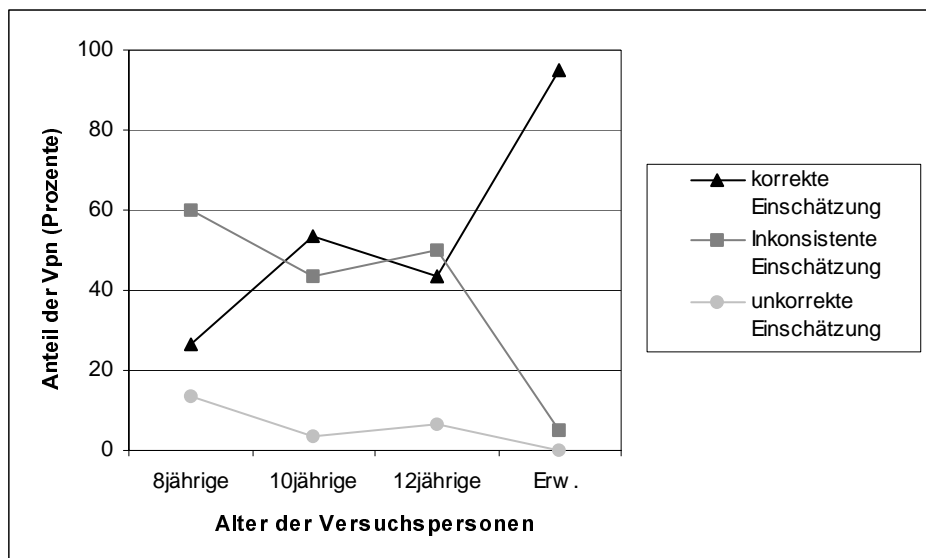


Abbildung 3.3.23. Konsistent korrekte und unkorrekte sowie inkonsistente Einschätzungen im Hinblick auf die prädiktive Regelanwendung anhand der subjektiven Sicherheit bei der unterschiedlichen Intensitätskombination in Prozenten pro Altersgruppe.

Die Überprüfung der Verteilungen der beiden Variablen auf Unabhängigkeit mit dem χ^2 -Test nach Pearson zeigte einen höchst signifikanten Wert, $\chi^2(6, N = 110) = 24.6$; $p < .001$. Da aber vier Zellen (33.3%) eine erwartete Häufigkeit < 5 aufwiesen, wurde ein erneuter Test ohne die dritte Kategorie der konsistent unkorrekten Einschätzungen durchgeführt, $\chi^2(3, N = 103) = 19.86$; $p < .001$.

In Tabelle 3.3.11 sind die Häufigkeiten und prozentualen Anteile der verschiedenen Einschätzungen schematisch dargestellt.

Tabelle 3.3.11

Darstellung der Häufigkeiten und Prozente der konsistenten Einschätzungen bei der physikalisch unmöglichen und möglichen Mischungsart mit der Sicherheitsskala nach Beobachtung der Mischungsvorgänge bei der unterschiedlichen Intensitätskombination

Prädiktiver Regeltyp	Konsistenz der Urteile beim bewussten Umgang mit der Regel					
	korrekt		inkonsistent		unkorrekt	
Durchschnitt	52	(65.8%)	25	(31.6%)	2	(2.5%)
Inkonsistenz	2	(13.3%)	13	(86.7%)	0	(0%)
Addition	2	(12.5%)	9	(56.3%)	5	(31.1%)

Anmerkung. In der linken Spalte sind die Häufigkeiten der Versuchspersonen und in der rechten Spalte die prozentualen Anteile bezogen auf den prädiktiven Regeltyp in Klammern aufgeführt.

Um eine Aussage über den Entwicklungsverlauf des konsistenten Umgangs mit der prädiktiven Regel nach Beobachtung der Mischungsvorgänge machen zu können, wurden die Häufigkeiten konsistenter korrekter und unkorrekter sowie inkonsistenter Einschätzungen pro Altersgruppe in Tabelle 3.3.12 aufgeführt.

Tabelle 3.3.12

Häufigkeiten konsistenter korrekter und unkorrekter sowie indifferenter Einschätzungen mit der Sicherheitsskala nach Beobachtung der Mischungsvorgänge pro Altersgruppe und Art der prädiktiven Regelanwendung bei der unterschiedlichen Intensitätskombination

Altersgruppe	Prädiktive Regelanwendung								
	Durchschnitt			Inkonsistent			Addition		
	korrekt	inkons.	unkorr.	korrekt	inkons.	unkorr.	korrekt	inkons.	unkorr.
Achtjährige (n=30)	7	7	1	1	7			4	3
Zehnjährige (n=30)	14	8	1	1	2		1	3	
Zwölfjährige (n=30)	13	10			4			1	2
Erwachsene (n=20)	18						1	1	

Die Verteilungen der Häufigkeiten bei der Durchschnittsregel auf die verschiedenen Altersgruppen zeigen, dass die korrekte Einschätzung mit zunehmendem Alter anstieg. Ein Anstieg zeigte sich von den Acht- zu den Zehnjährigen und von den Zwölfjährigen zu den Erwachsenen. Nur zwei Personen liessen sich insgesamt konsistent umstimmen.

Wie schon bei der gleichen Intensitätskombination widerspiegeln sich bei den inkonsistenten Einschätzungen der prädiktiven Durchschnittsregelanwender die Verunsicherungen durch das physikalisch unmögliche Mischungsresultat (vgl. Abbildung 3.3.18): Die Acht- und Zwölfjährigen liessen sich eher verunsichern, was zu inkonsistenten Urteilen führte. Im Vergleich zur gleichen Intensitätskombination urteilten hier aber tendenziell mehr Durchschnittsregelanwender inkonsistent.

Bei den Verteilungen der prädiktiven Addierer pro Altersgruppe zeigte sich das gleiche Muster wie bei den einzelnen Reaktionen auf die Mischungsresultate: Bei den Acht- und Zwölfjährigen bestanden Tendenzen zur Stabilität und bei den Zehnjährigen und den Erwachsenen Tendenzen zu inkonsistenten Einschätzungen.

Vergleich der korrekten Einschätzungen der abhängigen Variablen „prädiktive Regelanwendung“, „Diskriminierungsfähigkeit“ und „subjektive Sicherheit“

Im Anschluss an die detaillierten Auswertungen der einzelnen abhängigen Variablen werden diese noch miteinander verglichen. Die korrekten Einschätzungen der abhängigen Variablen „prädiktive Regelanwendung“, „Diskriminierungsfähigkeit“ und „subjektive Sicherheit“ können im Hinblick auf den unterschiedlichen Zugang zum physikalischen Mischungsgesetz miteinander in Beziehung gebracht werden. Angesichts der unterschiedlichen Behandlung der beiden Intensitätskombinationen erscheint es sinnvoll, diese getrennt darzustellen.

Die Verteilungen auf je zwei Variablen konnten mit einer 2x2-Kreuztabelle dargestellt werden. Ein allfälliger Zusammenhang zwischen den Variablen wurde mit dem Vierfelder- χ^2 -Test nach Pearson überprüft. Da bei der Einschätzungen der subjektiven Sicherheit nach Beobachtung des Mischungsvorganges keine Urteile von sechsjährigen Kindern vorlagen, beschränkte sich der Vergleich mit dieser abhängigen Variablen stets auf die Altersgruppen der acht- bis zwölfjährigen Kinder sowie den Erwachsenen.

Gleiche Intensitätskombination

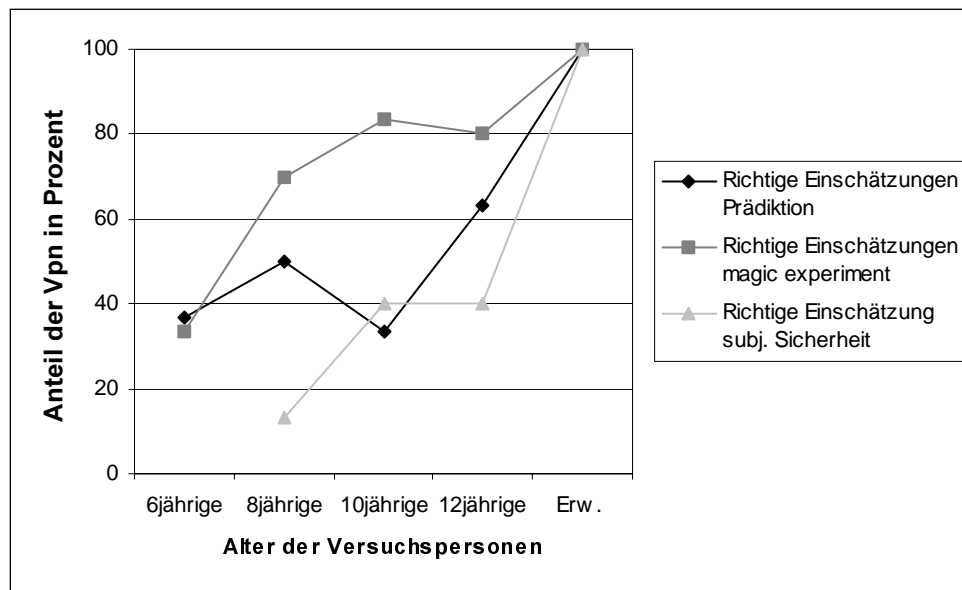


Abbildung 3.3.24. Prozentualer Anteil korrekter Einschätzungen bei den drei abhängigen Variablen „prädiktive Regelanwendung“, „Diskriminierungsfähigkeit“ und „subjektive Sicherheit“ pro Altersgruppe bei der gleichen Intensitätskombination.

Der Vierfelder- χ^2 -Test zeigte einen sehr signifikanten Unterschied der Verteilungen richtiger und falscher Einschätzungen zwischen der Prädiktionsbedingung und dem *magic experiment*, $\chi^2(1, N = 140) = 12.51, p < .001$. Der Vierfelder- χ^2 -Test wies auch einen signifikanten Unterschied der Verteilungen richtiger und falscher Einschätzungen zwischen der Prädiktionsbedingung und der subjektiven Sicherheit auf, $\chi^2(1, N = 110) = 44.28, p < .001$. Zudem unterschieden sich die Verteilungen richtiger und falscher Einschätzungen zwischen dem *magic experiment* und der subjektiven Sicherheit gemäss dem Vierfelder- χ^2 -Test höchst signifikant, $\chi^2(1, N = 110) = 14.84, p < .001$.

Unterschiedliche Intensitätskombination

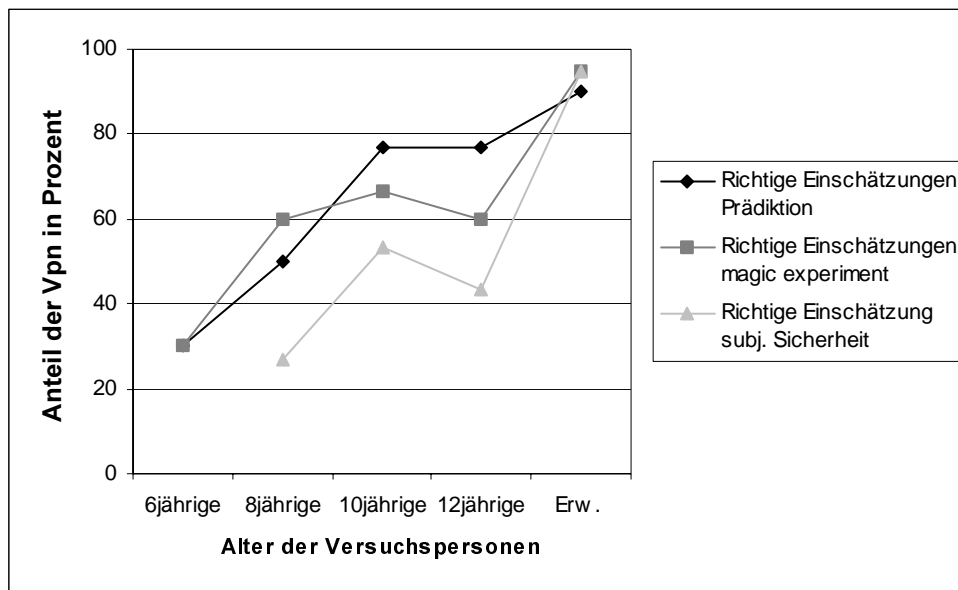


Abbildung 3.3.25. Prozentualer Anteil korrekter Einschätzungen bei den drei abhängigen Variablen „prädiktive Regelanwendung“, „Diskriminierungsfähigkeit“ und „subjektive Sicherheit“ pro Altersgruppe bei der gleichen Intensitätskombination.

Auch bei der unterschiedlichen Intensitätskombination zeigte der Vierfelder- χ^2 -Test einen höchst signifikanten Unterschied der Verteilungen richtiger und falscher Einschätzungen zwischen der Prädiktionsbedingung und dem *magic experiment*, $\chi^2(1, N = 140) = 15.99, p < .001$. Mit dem gleichen Testverfahren wurden auch die Verteilungen richtiger und falscher Einschätzungen zwischen der Prädiktionsbedingung und der subjektiven Sicherheit überprüft, $\chi^2(1, N = 110) = 24.95, p < .001$. Die Überprüfung der Verteilungen richtiger und falscher Einschätzungen zwischen dem *magic experiment* und der subjektiven Sicherheit mit dem Vierfelder- χ^2 -Test schliesslich wies ebenfalls auf eine sehr deutliche Abhängigkeit der beiden Variablen hin, $\chi^2(1, N = 110) = 7.79, p < .01$.

Inhaltliche Validität der Zauberskala

Die Frage nach der inhaltlichen Validität der Zauberskala kann von zwei Seiten angegangen werden.

1. Konnten die Kinder die Skala überhaupt inhaltlich korrekt anwenden?
2. Was heisst überhaupt fest gezaubert?

Zur Beantwortung der ersten Frage geben am ehesten die Entwicklungsverläufe der Durchschnittsregelanwender, welche auch richtig diskriminierten Auskunft, da von ihnen am ehesten erwartet werden konnte, dass sie nach einer Regel urteilten. Die Entwicklungsverläufe dieser Versuchspersonen sind in Abbildung 3.3.26 anhand von Mittelwerten und Standardfehlern pro Altersgruppe für alle vier Versuchsdurchgänge dargestellt.

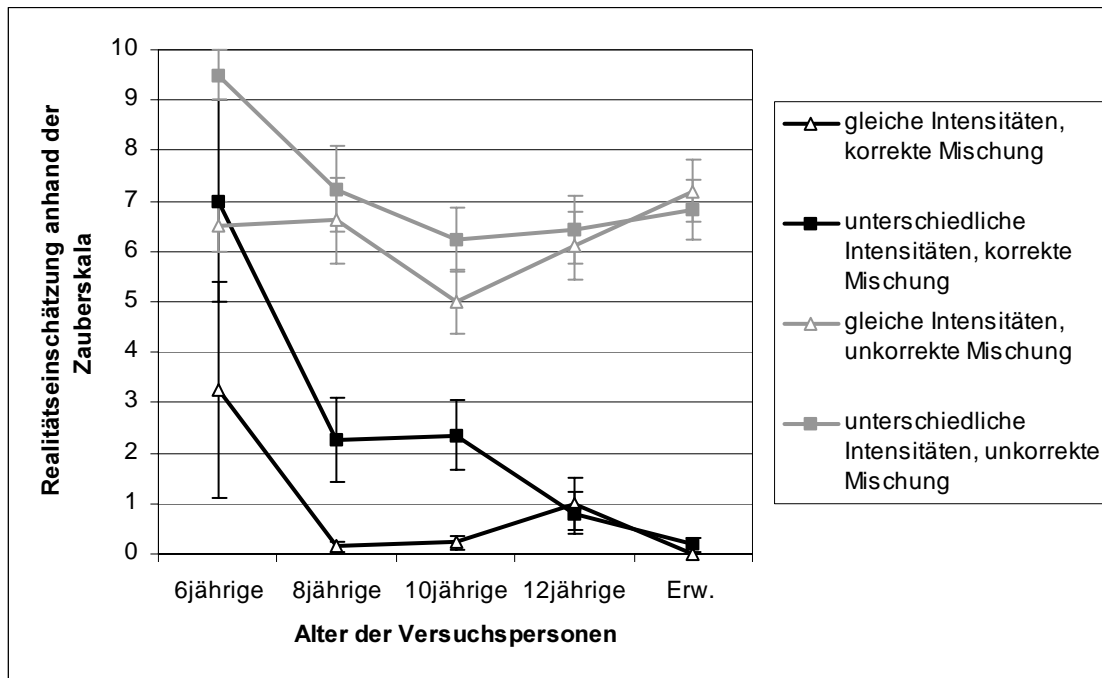


Abbildung 3.3.26. Mittelwerte und Standardfehler der Einschätzungen auf der Zauberskala pro Altersgruppe und Versuchsdurchgang bei den Durchschnittsregelanwendern aus Phase 1, welche die Mischungsergebnisse im *magic experiment* korrekt diskriminiert hatten (Gleiche Intensitätskombination: $n = 63$; unterschiedliche Intensitätskombination: $n = 64$).

In Abbildung 3.3.26 zeigt sich bei der gleichen Intensitätskombination, dass die physikalisch möglichen Mischungsergebnisse als fast nicht gezaubert eingeschätzt wurden mit Ausnahme der Sechsjährigen. Es fragt sich demnach nur bei den Sechsjährigen, ob die Zauberskala möglicherweise nicht korrekt angewendet werden konnte. Deshalb wurden die Daten dieser vier Kinder genauer analysiert. Da alle vier Sechsjährigen die Mischungsergebnisse auf der Wahrnehmungsebene richtig eingeschätzt hatten, konnte davon ausgegangen werden, dass sie bemerkt hatten, dass bei der Mischung von zwei gleich hellen Flüssigkeiten, die Farbintensität gleich geblieben war. Zwei dieser vier Kinder hatten bei den physikalisch möglichen Mischungen einen Wert von null auf der Zauberskala eingeschätzt. Auch die Urteile dieser beiden Kinder in den anderen Versuchsb

dingungen liessen auf ein adäquates Verständnis der Zauberskala schliessen. Die anderen beiden Sechsjährigen allerdings schätzten auch die physikalisch möglichen Mischungsresultate als gezaubert (Wert 4 bzw. 9) ein. Beide schätzten die physikalisch unmögliche Mischung aber mit maximalen 10 Zauberwerten ein. Demnach unterschied ein Kind die beiden Resultate doch mit 6 Zauberwerten, das andere aber nur mit einem Wert. Allerdings schätzten beide Kinder auch den Wassertrick als hoch bzw. maximal gezaubert ein. Diese Analyse spricht nicht gegen ein Verständnis der Zauberskala, es stellt sich vielmehr die Frage, wieso auch die physikalisch möglichen Mischungen nicht eindeutig als nicht-gezaubert taxiert werden konnten. Es wird vermutet, dass den Kindern der Mut oder vielleicht auch die Gewissheit fehlte, die physikalisch korrekten Mischungsresultate als nicht-gezaubert einzuschätzen.

Die Verteilungen des Ausmasses der Diskrepanz bzw. Diskriminierung zwischen der physikalisch unmöglichen und der möglichen Mischungsart der gleichen Intensitätskombination zeigten einerseits eine Normalverteilung mit einem Modalwert von 5 und andererseits eine hohe Anzahl von maximaler Unterscheidung von 10. Die Zauberskala wurde bei den physikalisch unmöglichen Mischungsresultaten auf zwei verschiedene Weisen angewendet: Das Mischungsresultat war generell maximal gezaubert oder es wurde eine Abstufung – vermutlich im Vergleich mit den anderen Versuchsdurchgängen und dem Probeversuchsdurchgang – vorgenommen. Die gleichen Antwortmuster waren auch bei der unterschiedlichen Intensitätskombination festzustellen. Die Verteilungen der Diskriminierungseinschätzungen zeigten eine Normalverteilung mit einem Modalwert von 4 und einigen Einschätzungen bei der Maximaldiskrepanz von 10.

Interessant ist auch der Vergleich zu den Verteilungen bei den Additionsregelanwendern. Bei ihnen zeigte sich bei beiden Intensitätskombinationen eine Tendenz zu Diskriminierungen in geringerem Ausmass: Von insgesamt 9 bzw. 10 Addierern, die auch unkorrekt diskriminierten, unterschieden je 3 die Mischungsresultate mit -1 und -2 ein. Die geringen Unterschiede zwischen den beiden Mischungsresultaten bei den Additionsregelanwendern weisen vermutlich eher auf eine Unsicherheit bei der Einschätzung hin als auf ein inhaltliches Missverständnis der Skala.

Es bleibt noch die Frage nach den Einschätzungen bei der unterschiedlichen Intensitätskombination zu beantworten. Abbildung 3.3.26 zeigt, dass die physikalisch korrekten

Mischungsergebnisse bei der unterschiedlichen Intensitätskombination von den sechs- bis zehnjährigen Kindern als mehr gezaubert eingeschätzt wurden als diejenigen bei der gleichen Intensitätskombination. Eine Vermutung, die allerdings durch das Gesamtergebnis des *magic experiment* gestützt wird, liegt darin, dass die Mischungen unterschiedlicher Intensitätskombinationen generell schwieriger einzuschätzen waren und zu mehr Unsicherheit geführt hatten.

Zur Aufklärung des inhaltlichen Verständnisses der Zauberskala bei der unterschiedlichen Intensitätskombination wird hier der Frage nachgegangen, wie bei korrekten Regelanwendern indifferente Einschätzungen (d. h. beide Mischungsergebnisse gleich) zu Stande kamen. Hierzu wurden nur diejenigen prädiktiven Durchschnittsregelanwender ausgewählt, welche auch bei der subjektiven Sicherheitseinschätzung konsistent richtig geurteilt hatten. Insgesamt 7 Versuchspersonen hatten trotz richtiger Regel die beiden Mischungsergebnisse auf der Zauberskala nicht diskriminiert und hatten Werte > 0 auf der Zauberskala angegeben. Diese Gruppe setzte sich aus 3 Acht-, 3 Zehnjährigen und 1 Erwachsene zusammen. Vier dieser Versuchspersonen zeigten bei der gleichen Intensitätskombination ein identisches Urteilmuster, welches auf ein inhaltliches Verständnis der Zauberskala schließen lässt: Sie waren alle Durchschnittsregelanwender und schätzten die physikalischen möglichen Mischungsergebnisse als nicht gezaubert ein. Der Wassertrick wurde zudem bei allen vier Personen mit sehr hohen Werten auf der Zauberskala beurteilt. Bei drei der vier Personen war festzustellen, dass sie eines der Mischungsergebnisse auf der Wahrnehmungsebene unkorrekt beurteilt hatten. Dadurch wird verständlich, weshalb sie die beiden Mischungsarten nicht differenzieren konnten. Bei den drei anderen der sieben indifferenten Versuchspersonen war bei zwei Personen der Wassertrick ebenfalls als sehr hoch gezaubert eingeschätzt worden. Die letzte Person, deren indifferentes Urteilsverhalten mit der Zauberskala noch aufzuklären bleibt, hatte auch bei der gleichen Intensitätskombination indifferent geurteilt, hier aber auch auf der Wahrnehmungsebene ein Mischungsergebnis falsch eingeschätzt. Da aber dieser zehnjährige Junge beide Mischungsergebnisse der unterschiedlichen Intensitätskombination auf der Wahrnehmungsebene richtig beurteilt hatte, lässt sich sein Verhalten bei der Einschätzung auf der Zauberskala nicht aufgrund der Daten erklären.

Die Beantwortung der zweiten Eingangsfrage zur inhaltlichen Validität der Zauberskala richtet sich auf allfällige Differenzierungen verschiedener Farbtricks anhand der Zauberskala. Es stellen sich also die Fragen: War der Wassertrick mehr gezaubert als der

Farbmischungstrick? War der Farbmischungstrick bei der gleichen Intensitätskombination (Differenz auf der Farbskala = 2) mehr gezaubert als derjenige bei der unterschiedlichen Intensitätskombination (Differenz auf der Farbskala = 1)? Der Unterschied zwischen den beiden Farbmischungstricks und dem Wassertrick wurde in Abbildung 3.3.9 deutlich: Der Wassertrick war in allen Altersgruppen mehr gezaubert. Zwischen den beiden Farbmischungstricks zeigte sich bei den generellen Einschätzungen (siehe Abbildung 3.3.2) aller Versuchspersonen kein Unterschied. Die Einschätzungen der Durchschnittsregelanwender, welche auch korrekt diskriminierten wiesen fast durchgängig höhere Werte bei der physikalisch unkorrekten Mischung unterschiedlicher Intensitäten auf. Allerdings waren die Mittelwerte der Diskriminierungen bei der gleichen Intensitätskombinationen (Mittelwert: 5.89; Modalwert: 5) höher als bei der unterschiedlichen Intensitätskombination (Mittelwert: 5.18; Modalwert: 4), so dass kein eindeutiger Schluss gezogen werden kann.

3.3.5 Diskussion

Angesichts der Fülle der Resultate der Phase 2 der vorliegenden Untersuchung scheint es sinnvoll, die Ergebnisse zunächst getrennt für die einzelnen abhängigen Variablen zusammenzufassen und im Hinblick auf die spezifischen Fragestellungen zu diskutieren. Erst im Anschluss daran werden die Ergebnisse insgesamt zusammengefasst und im Zusammenhang mit den Ausgangshypothesen diskutiert.

Realitätseinschätzungen anhand der Zauberskala

Allgemeine Aussagen zu den Realitätseinschätzungen

Abbildung 3.3.2 (siehe S. 109) und nachfolgende Berechnungen zeigten, dass die beiden Mischungsarten, d. h. die physikalisch korrekten und unkorrekten Mischungsergebnisse, bei der Einschätzung mit der Zauberskala ab dem Alter von acht Jahren höchst signifikant unterschieden wurden. Die Unterscheidung zwischen den physikalisch korrekten und unkorrekten Mischungsergebnissen hinsichtlich ihres Realitätscharakters gelang insgesamt sehr gut. Es zeigte sich auch, dass diese Unterscheidung bei den Erwachsenen viel deutlicher war als bei den Kindern. Zudem wirkten sich die Mischungsergebnisse

bei den Sechsjährigen bei beiden Intensitätsbedingungen nicht auf die Einschätzung mit der Zauberskala aus.

Hinsichtlich des Effektes der Intensitätskombination auf die Realitätseinschätzung zeigten sich deutliche Unterschiede bei allen Kindern zwischen den Mischungen gleicher und unterschiedlicher Intensitäten. Diese waren allerdings nur bei den physikalisch korrekten Mischungsresultaten auszumachen: Die Kinder schätzten die physikalisch korrekten Mischungen unterschiedlicher Ursprungsintensitäten als mehr gezaubert ein als diejenigen gleicher Intensitäten. Die physikalisch korrekten Mischungen unterschiedlicher Intensitäten waren also für die Kinder schwieriger als solche zu erkennen. Da diese aber bei der Einschätzung auf der Wahrnehmungsebene deutlich realistischer eingeschätzt wurden als die korrekten Resultate bei der Mischung gleicher Intensitäten, war der Unterschied nicht auf eine allfällige Schwierigkeit auf der Wahrnehmungsebene zurückzuführen. Bei den physikalisch unkorrekten Mischungsarten hingegen bestand kein Unterschied bei der Realitätseinschätzung der Mischungsresultate verschiedener Intensitätskombinationen.

Dieses Resultat spricht dafür, dass die Manipulationen der beiden Intensitätskombinationen gelungen waren. Obwohl die physikalisch unkorrekte Mischung gleicher Intensitäten mit zwei Abständen der Farbskala vom korrekten Mischungsresultat abwich, wurde sie nicht von der Mischung von unterschiedlichen Intensitäten, welche nur mit einem Abstand abwich, unterschieden. Die beiden physikalisch unkorrekten Resultate der Farbmischungen wurden aber deutlich – auch von den Sechsjährigen – als weniger gezaubert eingeschätzt als die Mischung beim Probeversuchsdurchgang. Dieses Resultat wird im Hinblick auf die Validität der Zauberskala als bedeutend betrachtet.

Unterschiede bei der Realitätseinschätzung der Mischungsresultate bei verschiedenen Regelanwendern

Ganz allgemein kann ausgesagt werden, dass sich die prädiktive Regelanwendung im Sinne der Hypothese auf die Einschätzungen der beiden Mischungsarten auswirkte. Die korrekten Durchschnittsregelanwender waren bei der Einschätzung auf der Zauberskala bei beiden Intensitätskombinationen bei den physikalisch korrekten am tiefsten und bei den unkorrekten am höchsten. Die Durchschnittsbilder aus der prädiktiven Aufgabe urteilten beim *magic experiment* mehrheitlich ihrer Erwartung entsprechend. Anders sah

es bei den Addierern aus der prädiktiven Aufgabe aus: Da auch sie die physikalisch korrekten Mischungsresultate in beiden Bedingungen insgesamt auf der Zauberskala tiefer einschätzten als die unkorrekten, urteilten sie im Mittel nicht entsprechend ihrer früheren prädiktiven Aussage. Es kann eindeutig gesagt werden, dass die Durchschnittsbilder bei beiden Intensitätsbedingungen die beiden Mischungsresultate unabhängig vom Vorzeichen viel deutlicher unterschieden als die Addierer und die inkonsistenten Regelanwender. Während die Addierer die Mischungsresultate bei der Kombination gleicher Ursprungsintensitäten tendenziell deutlicher unterschieden als bei der Kombination unterschiedlicher Intensitäten, geschah dies bei den inkonsistenten Regelanwendern in statistisch relevantem Ausmass. Die Unterschiede der verschiedenen Regelanwender bei den Einschätzungen mit der Zauberskala werden erst im Abschnitt zur Diskrimination im Hinblick auf die Stabilitätshypothese diskutiert.

Unterschiede bei der Realitätseinschätzung der Mischungsresultate bei verschiedenen Regelanwendern im Entwicklungsverlauf

Entwicklung bei den Durchschnittsregelanwendern

Bei den Mittelwerten der Durchschnittsbilder aus Phase 1 zeigte sich, dass die physikalisch unkorrekten Mischungsarten bei beiden Intensitätskombinationen bei allen Altersgruppen ähnlich hoch eingeschätzt wurden. Bei der Beurteilung der physikalisch möglichen Mischungsresultate war bei beiden Intensitätskombinationen ein deutlicher Entwicklungsverlauf hin zu sehr tiefen Einschätzungen auf der Zauberskala zu beobachten. Allerdings war der Entwicklungsverlauf bei der unterschiedlichen Intensitätskombination leicht verzögert. Aus der Auswertung und Diskussion zur Validität der Zauberskala lassen die erhöhten Zauberwerte bei den physikalisch möglichen Mischungen bei den jüngeren Kindern vermutlich auf eine Unsicherheit bei der Einschätzung der Mischungsresultate schliessen. Gemäss dieser Vermutung wäre der verzögerte Entwicklungsverlauf auf eine geringere Stabilität der Durchschnittsregel bei der unterschiedlichen Intensitätskombination zurückzuführen.

Entwicklung bei den Additionsregelanwendern

Grundsätzlich sind die Ergebnisse bei den Additionsregelanwendern statistisch weniger gut abgesichert, weil die Gruppe der Addierer viel kleiner war als diejenige der Durch-

schnittsregelanwender. Die Schlussfolgerungen auf der Basis dieser Resultate sind deshalb spekulativer Art.

Die Reaktionen auf die physikalisch unmöglichen Mischungsresultate, welche den Erwartungen der Addierer aus Phase 1 hypothetisch entsprochen hatten, waren bei beiden Intensitätskombinationen ähnlich: Die unkorrekten Mischungsresultate waren für zunehmend ältere Kinder weniger gezaubert. Je älter die Kinder waren, umso eher wurde die Regel beim erwarteten Mischungsresultat beibehalten.

Bei der physikalisch korrekten Mischungsart reagierten die Addierer auf eine Weise, die schwieriger zu interpretieren ist. Auffällig waren die Achtjährigen bei beiden Intensitätskombinationen und die Zwölfjährigen bei der gleichen Intensitätskombination, da sie die physikalisch möglichen Mischungen als ähnlich wenig gezaubert einstufen wie die Durchschnittsregelanwender. Nur die Zehnjährigen – abgesehen von den Sechsjährigen – gaben bei beiden Intensitätskombinationen und die Zwölfjährigen bei der unterschiedlichen Mischungsart höhere Zauberverte. Hinsichtlich der Stabilitätshypothese zeigte sich demnach ein kompliziertes Bild: Die Achtjährigen waren bei beiden Intensitätskombinationen instabil und die Zehnjährigen am stabilsten. Die Zwölfjährigen blieben ihrer Regel bei der unterschiedlichen Intensitätskombination deutlich treuer als bei der gleichen Intensitätskombination.

Die Entwicklungsverläufe der inkonsistenten Regelanwender werden erst im nächsten Abschnitt zur Diskrimination zusammenfassend betrachtet und diskutiert.

Diskrimination zwischen den beiden Mischungsarten

Das Ausmass der Diskrimination zwischen den beiden Mischungsresultaten wurde berechnet durch die Differenz zwischen den Einschätzungen auf der Zauberskala des physikalisch unkorrekten Mischungsresultates und des physikalisch korrekten Mischungsresultates. Positive Diskriminierungen entsprachen höheren Werten auf der Zauberskala bei den physikalisch unkorrekten Mischungen als bei den korrekten Mischungen und somit korrekten Realitätseinschätzungen.

Gemäss den Ausgangshypothesen kann das Ausmass der Diskrimination zwischen den beiden Mischungsarten im Hinblick auf zwei verschiedene theoretische Ansätze interpretiert und diskutiert werden:

1. Ausmass der Unterscheidung zwischen den beiden Mischungsresultaten durch Einschätzung des Realitätscharakters als Ausdruck einer Mischungsregel. Vergleich der Bedingung des *magic experiment* mit der prädiktiven Bedingung: Unter welcher Bedingung wurden die Mischungsresultate besser beurteilt?
2. Ausmass und Vorzeichen der Diskrimination in Abhängigkeit von der prädiktiven Regelanwendung: Je diskrepanter die beiden Einschätzungen (Vorzeichen!) umso deutlicher wurde nach einer Regel geurteilt und umso stabiler bzw. instabiler die Regel (Vorzeichen!).

Generelle Diskriminationsfähigkeit

Es zeigte sich, dass die Mischungsresultate der beiden Intensitätskombinationen unterschiedlich gut diskriminiert wurden: Die Mischungsresultate der Mischung von zwei unterschiedlich hellen Flüssigkeiten wurden signifikant schlechter diskriminiert, d. h. sie waren schwieriger einzuschätzen. Wie die Auswertung der einzelnen Einschätzungen der Mischungsresultate gezeigt hat, liegt der Unterschied in der Einschätzung der physikalisch korrekten Mischungsart begründet. Das physikalisch korrekte Mischungsresultat der gleichen Intensitätskombination wurde als weniger gezaubert eingeschätzt als dasjenige der unterschiedlichen Intensitätskombination. Mögliche Gründe für diesen Unterschied werden im Abschnitt zur Betrachtung der Ergebnisse im Lichte des Ansatzes des adaptiven Denkens (S. 162ff.) diskutiert.

Die Diskriminationsfähigkeit wurde insgesamt im Entwicklungsverlauf signifikant besser. Allerdings stieg das Ausmass der Diskrimination nicht mit zunehmendem Alter stetig an, sondern es zeigte sich bei beiden Intensitätskombinationen annäherungsweise ein Plateau bei den Schulkindern. Während bei der unterschiedlichen Intensitätskombination die Häufigkeit korrekter Diskriminationen bei den drei Altersgruppen der Schulkin- der nicht anstieg, nahm sie bei der gleichen Intensitätskombination zwischen den Acht- und Zehnjährigen doch noch zu. Die Hypothese, dass die Realitätseinschätzungen mit zunehmendem Alter besser gelingen würden, war damit bestätigt.

Diskriminationsunterschiede verschiedener Regelanwender

Die Ergebnisse hinsichtlich der Diskriminationen der beiden Mischungsarten wiesen signifikante Unterschiede zwischen den prädiktiven Durchschnitts- und Additions-

regelanwender auf. Die Durchschnittsregelanwender diskriminierten bei beiden Intensitätskombinationen am höchsten. Die Addierer diskriminierten bei den Mischungsergebnissen der gleichen Intensitätskombination am tiefsten, während die prädiktiv inkonsistenten Regelanwender zwischen den beiden anderen lagen. Bei der Diskrimination der unterschiedlichen Intensitätskombination waren die Inkonsistenten allerdings noch tiefer als die Addierer.

Interessant ist an dieser Stelle auch der Vergleich zwischen der prädiktiven Voraussage des Mischungsergebnisses und der kategorisierten Einschätzung beim *magic experiment*. Gab es eine signifikante Verbesserung der falschen Mischungsregeln durch Beobachtung der Mischung? Es zeigte sich, dass die Mehrheit der prädiktiven Additions- und inkonsistenten Regelanwender die Mischungsergebnisse der gleichen Intensitätskombination korrekt diskriminierten. Hier fand also eine deutliche Verbesserung im Vergleich zur prädiktiven Mischungsaufgabe statt. Bei der unterschiedlichen Intensitätskombination sah es aber anders aus: Die Mehrheit der inkonsistenten Regelanwender urteilte unkorrekt und bei den Additionsregelanwendern hielten sich die korrekten und unkorrekten Einschätzer die Waage.

Hinsichtlich der Stabilitätshypothese lässt sich sagen, dass die Durchschnittsregelanwender in hohem Ausmass stabil blieben (84 % bei der gleichen Intensitätskombination, 73% bei der unterschiedlichen Intensitätskombination). Bei den Additionsregelanwendern war bei der unterschiedlichen Intensitätskombination eine höhere Stabilität (42%) zu verzeichnen als bei der gleichen Intensitätskombination (33%). Damit erwies sich das „richtige Wissen“ im *magic experiment* als viel stabiler als das „falsche“ Wissen.

Diskriminationsunterschiede verschiedener Regelanwender im Entwicklungsverlauf

Entwicklung bei den Durchschnittsregelanwendern

Die Durchschnittsregelanwender aus Phase 1 diskriminierten bei beiden Intensitätskombinationen ab dem Alter von 8 Jahren vorwiegend korrekt. Ein grosser Entwicklungssprung scheint sich zwischen dem Alter von 6 und 8 Jahren zu vollziehen. Allerdings diskriminierten die Erwachsenen nochmals deutlicher als die Zwölfjährigen. Die Diskriminierung der beiden Mischungsarten fiel den Durchschnittsregelanwendern bei der unterschiedlichen Intensitätskombination offenbar schwerer als bei der gleichen In-

tensitätskombination, was sich durch häufigere indifferente Einschätzungen in der Bedingung unterschiedlicher Intensitäten äusserte.

Bezüglich einer Aussage über die Stabilität der Durchschnittsregel eignen sich die Häufigkeiten korrekter, indifferenter und unkorrekter Diskriminationen über die verschiedenen Altersgruppen besser als der Entwicklungsverlauf der Mittelwerte. Es zeigte sich insbesondere ein Stabilitätsunterschied zwischen den Sechsjährigen und den übrigen Versuchspersonen. Bei beiden Intensitätskombinationen hatte die Mehrheit der Sechsjährigen die beiden Mischungsresultate nicht gemäss der Durchschnittsregel unterschieden. Ab dem Alter von 8 Jahren diskriminierten die prädiktiven Durchschnittsregelanwender bei beiden Intensitätskombinationen vorwiegend korrekt. Allerdings waren bei der unterschiedlichen Intensitätskombination mehr indifferente Einschätzungen zu verzeichnen. Die bei der Einschätzung des physikalisch möglichen Mischungsresultates mit der Zauberskala beobachteten Unsicherheiten bei der unterschiedlichen Intensitätskombination äusserten sich demnach bei der Diskriminierungsfähigkeit in vermehrter Indifferenz. Diese nahm im Entwicklungsverlauf sogar tendenziell zu.

Entwicklung bei den Additionsregelanwendern

Bei den Addierern aus Phase 1 war im Entwicklungsverlauf der Kinder bei der gleichen Intensitätskombination kaum eine Verbesserung der Diskrimination zu beobachten. Bei der unterschiedlichen Intensitätskombination war sogar eher ein Trend Richtung verschlechterter Diskriminierungsfähigkeit festzustellen. Allerdings diskriminierten hier die beiden Erwachsenen wieder korrekt.

Bezüglich einer Aussage über die Stabilität der prädiktiven Additionsregel wurden die Häufigkeiten korrekter, indifferenter und unkorrekter Diskriminationen herangezogen. Obwohl die Mittelwertsdarstellung bei den Achtjährigen auf eine höhere Diskriminationsfähigkeit gegenüber den älteren Kindern hinwies, zeigte sich diese Tendenz bei den Häufigkeiten korrekter, indifferenter und unkorrekter Einschätzungen bei der gleichen Intensitätskombination nicht mehr. Die Addierer urteilten mit leichter Tendenz mit zunehmendem Alter entgegen ihrer prädiktiven Regel. Die Anzahl derjenigen, welche ihrer Regel treu blieben, war mit durchschnittlich zwei Kindern in den verschiedenen Altersgruppen ähnlich hoch. Daraus lässt sich schliessen, dass die Achtjährigen zwar nicht häufiger, dafür aber in höherem Ausmass anhand der Zauberskala diskriminierten als

die älteren Kinder. Bei der unterschiedlichen Intensitätskombination wurden die Kinder, welche die beiden Mischungsresultate korrekt diskriminierten, mit zunehmendem Alter aber seltener. Aufgrund der kleinen Stichprobengrößen fallen generelle Aussagen schwer, dennoch wurden folgende Trends formuliert: Bei der gleichen Intensitätskombination wurden die Additionsregelanwender mit zunehmendem Alter instabiler und bei der unterschiedlichen Intensitätskombination etwas stabiler.

Entwicklung bei den inkonsistenten Regelanwendern

Bei den prädiktiven inkonsistenten Regelanwendern war bei der gleichen Intensitätskombination eine interessante eindeutige Entwicklung zu beobachten: Mit zunehmendem Alter wurden die beiden Mischungsarten deutlicher diskriminiert. Die zehn- und zwölfjährigen inkonsistenten Regelanwender diskriminierten die beiden Mischungsresultate der gleichen Intensitätskombination fast genau so gut wie die gleichaltrigen Durchschnittsregelanwender. Die Analyse der einzelnen Einschätzungen der beiden Mischungsarten ergab, dass die zunehmend bessere Diskriminationsfähigkeit auf die Einschätzungen bei den physikalisch unmöglichen Mischungsresultaten zurückgeführt werden kann. Bei der unterschiedlichen Intensitätskombination fiel in erster Linie auf, dass die inkonsistenten Regelanwender aus der Phase 1 im *magic experiment* viel schlechter urteilten als bei der gleichen Intensitätskombination. Es zeigte sich eher ein Trend Richtung Verschlechterung der Diskrimination mit zunehmendem Alter. Die Einschätzungen bei den einzelnen Mischungsarten zeigten, dass diese Verschlechterung bei den älteren Kindern auf Schwierigkeiten bei der Einschätzung der physikalisch unkorrekten Mischungsart beruhte. Der Entwicklungsverlauf war ähnlich wie bei den Additionsregelanwendern: Das physikalisch unkorrekte Mischungsresultat wurde mit zunehmendem Alter als weniger gezaubert eingeschätzt.

Die mit dem Alter zunehmend bessere Diskrimination der inkonsistenten Regelanwender bei der gleichen Intensitätskombination ist aus zwei Gründen ein interessantes Resultat: Einerseits lässt es die Vermutung aufkommen, dass die Aufgabenpräsentation im *magic experiment* für die Einschätzung des Mischungsresultates gleich heller Flüssigkeiten zur richtigen Aufgabenlösung verleitete. Andererseits zeigte sich diese Verbesserung insbesondere bei denjenigen Altersgruppen deutlich, die in der prädiktiven Bedingung Mühe bekundeten, eine konsistente Regelanwendung zu zeigen. Wurde in Phase 1

durch die Aufgabenpräsentation bei ihnen eine korrekte Lösungsstrategie verhindert, so war diese Blockierung im *magic experiment* aufgehoben.

Subjektive Sicherheit des prädiktiven Mischungsurteils nach Rückmeldung durch die verschiedenen Mischungsergebnisse

Allgemeine Aussagen

Bei den Einschätzungen der eigenen Sicherheit bezüglich der prädiktiven Regel war ein statistisch relevanter Entwicklungsanstieg mit zunehmendem Alter zu beobachten. Da ein Anstieg aber erst bei den Zwölfjährigen sichtbar wurde, scheint die subjektive Sicherheit bzw. der bewusste Umgang mit der Regel im Vergleich zu den anderen abhängigen Variablen verzögert zu sein. Beide physikalisch korrekten Mischungsergebnisse vermittelten mehr Sicherheit als die unkorrekten. Bei den physikalisch unkorrekten Mischungsergebnissen verunsicherte allerdings das Resultat der unterschiedlichen Intensitätskombination viel stärker als dasjenige bei der gleichen Intensitätskombination. Dennoch wirkte sich beim bewussten Umgang mit der prädiktiven Regel die Intensitätskombination – im Vergleich mit den anderen abhängigen Variablen erstmals – nicht signifikant auf die Einschätzungen aus.

Die Zuweisungen der Urteile bei physikalisch korrekten und unkorrekten Mischungsergebnissen zu konsistent korrekten und unkorrekten sowie inkonsistenten Einschätzungen pro Intensitätskombination zeigten, dass vorwiegend konsistent korrekte und inkonsistente Einschätzungen gemacht wurden. Während die Achtjährigen mehrheitlich keinen konsistenten Umgang mit der Regel zeigten, hielten sich die konsistent korrekten und inkonsistenten Einschätzungen bei den Zehn- und Zwölfjährigen mit einem Anteil von gut 40 Prozent die Waage. Bei den Erwachsenen wurde fast nur noch ein konsistent korrekter Umgang mit der prädiktiven Regel beobachtet. Es scheint, als hätten die Achtjährigen mehr Mühe mit der Objektivierung ihrer Regeln.

Unterschiede im bewussten Umgang mit der prädiktiven Regel bei verschiedenen Regelanwendern

Ganz allgemein kann ausgesagt werden, dass sich die prädiktive Regelanwendung auf die subjektive Sicherheit der Regel nach der Beobachtung der Mischungsvorgänge bei beiden Mischungstypen auswirkte. Die korrekten Durchschnittsregel Anwender waren bei

der Einschätzung ihrer subjektiven Sicherheit der vorher angewendeten Regel bei beiden Intensitätskombinationen bei den physikalisch korrekten am höchsten und bei den unkorrekten tiefer aber immer noch recht hoch. Die Durchschnittsbilder aus der prädiktiven Aufgabe waren also beim *magic experiment* ziemlich stabil geblieben, obwohl die erwartete Mischung zu höherer Sicherheit führte als die unkorrekte. Anders sah es bei den Addierern aus der prädiktiven Aufgabe aus: Bei ihnen führte die Präsentation physikalisch korrekter Mischungsergebnisse in beiden Bedingungen insgesamt zu Verunsicherung, während die unkorrekten sie entsprechend ihrer früheren prädiktiven Aussage auch zu höherer Sicherheit führte. Obwohl die Sicherheitsunterschiede durch die verschiedenen Mischungsarten bei beiden Regelanwendern etwa ähnlich hoch waren, lagen die Addierer auf viel tieferem Sicherheitsniveau. Dies kann hinsichtlich der Stabilitätshypothese wie folgt interpretiert werden: Die Durchschnittsbilder waren in beiden Intensitätsbedingungen viel stabiler als die Addierer. Die Addierer zeigten zudem – wie schon bei der Zauberskala – bei der Kombination gleicher Ursprungsintensitäten eine geringere Stabilität als bei der Kombination unterschiedlicher Intensitäten.

Interpretiert man die Höhe des Sicherheitsniveaus als Ausdruck eines kognitiven Konflikts, so waren die Additionsregelanwender durch die Präsentation widersprüchlicher Ereignisse einem deutlich grösseren intraindividuellen Konflikt ausgesetzt.

Unterschiede im bewussten Umgang mit der prädiktiven Regeln bei verschiedenen Regelanwendern im Entwicklungsverlauf

Entwicklung bei den Durchschnittsregelanwendern

Bei beiden Intensitätskombinationen bestätigten die Durchschnittsregelanwender aus Phase 1 ihre Regel nach der Beobachtung der physikalisch korrekten Mischungsergebnisse in hohem Ausmass. Beiden Intensitätskombinationen gemeinsam war ausserdem, dass die korrekte Regel mit zunehmendem Alter deutlicher bestätigt wurde. Bei der Reaktion auf die physikalisch unkorrekten Mischungsarten zeigte sich bei den Kindern eine Verunsicherung ohne erkennbare Abnahme im Verlaufe der Entwicklung. Auffällig war hier insbesondere, dass die Achtjährigen durch die Rückmeldung eines unkorrekten Mischungsergebnisses im Mittel ihre korrekte Regel aus der Prädiktion sogar verworfen hatten.

Hinsichtlich eines konsistenten Umgangs mit der Regel zeigten sich ebenfalls bei den Achtjährigen Schwierigkeiten, insbesondere bei der gleichen Intensitätskombination, ihre korrekte Regel aus Phase 1 konsistent zu bestätigen. Allerdings war bei der gleichen Intensitätskombination eine leichte Verbesserung mit zunehmendem Alter festzustellen, die aber bei der unterschiedlichen Intensitätskombination etwas geringer war.

Entwicklung bei den Additionsregelanwendern

Aufgrund des uneinheitlichen Bildes werden hier die Entwicklungsverläufe der beiden Intensitätskombinationen zunächst getrennt betrachtet. Bei der gleichen Intensitätskombination stieg die Bestätigung der Additionsregel durch das physikalisch unkorrekte Mischungsresultat mit zunehmendem Alter an. Nach Beobachtung des physikalisch korrekten Mischungsresultates blieb die Reaktion bei allen Altersgruppen auf gleichem Niveau nahezu bei null, d. h. sie waren im Mittel ihrer Regel nicht mehr sicher. Insgesamt zeigte sich also eine leichte Tendenz zur Bestätigung der falschen Regel bei älteren Kindern bei der gleichen Intensitätskombination. Bei der unterschiedlichen Intensitätskombination waren die Entwicklungsverläufe bei beiden Mischungsarten sehr unstetig, aber durchaus kohärent. Während die Zehnjährigen und die Erwachsenen ihre falsche Regel verwarfen, waren die Acht- und Zwölfjährigen durch das physikalisch korrekte Mischungsresultat nur verunsichert und nach Rückmeldung des unkorrekten Mischungsresultates sogar recht hoch bestätigt. Insgesamt verwarfen also die Zehnjährigen im deutlichen Gegensatz zu den Acht- und Zwölfjährigen ihre Regel, während sie bei letzteren recht stabil war.

Hinsichtlich eines konsistenten Umgangs mit der Regel zeigte sich das gleiche Muster wie bei den Mittelwertsdarstellungen der einzelnen Versuchsdurchgänge. Bei der gleichen Intensitätskombination verwarf kein einziges Kind seine falsche Regel. Die Hälfte der Addierer urteilte inkonsistent und die andere Hälfte konsistent unkorrekt, d. h. sie blieben ihrer Regel treu. Bei den Zwölfjährigen war die Additionsregel sogar mehrheitlich stabil. Bei der unterschiedlichen Intensitätskombination zeigten sich Tendenzen zur Stabilität bei den Acht- und Zwölfjährigen sowie Tendenzen zur inkonsistenten Einschätzungen bei den Zehnjährigen und den Erwachsenen.

Entwicklung bei den inkonsistenten Regelanwendern

Die konsistenten Einschätzungen der subjektiven Regelsicherheit nach Beobachtung der Mischungsresultate bei den prädiktiv inkonsistenten Regelanwendern können hinsichtlich eines Lerneffektes beurteilt werden. Bei den Diskriminierungen der beiden Mischungsresultate wurde v. a. bei den älteren Kindern bei der gleichen Intensitätskombination die beiden Mischungsresultate eine gute Unterscheidungsfähigkeit beobachtet. Wie aber die konsistenten Einschätzungen mit der Sicherheitsskala im *magic experiment* zeigten, konnte diese Verbesserung von den inkonsistenten Regelanwendern nicht in einen bewussten Umgang mit der Mischungsregel umgesetzt werden. Die Mehrheit der vorher inkonsistenten Regelanwender blieb bei beiden Intensitätskombinationen inkonsistent.

Betrachtung der Ergebnisse im *magic experiment* im Lichte des Ansatzes des adaptiven Denkens

Die Einschätzungen der Mischungsresultate anhand ihres Realitätscharakters mit der Zauberskala konnten auch losgelöst von den Ergebnissen aus Phase 1 betrachtet und im Hinblick auf ihre Korrektheit ausgewertet werden. Dabei wurde das Verhältnis der Einschätzungen bei beiden Mischungsresultaten pro Intensitätskombination, bezeichnet als Diskriminierungsfähigkeit, als Mass für korrekte, unkorrekte bzw. indifferente Einschätzungen verwendet.

Es zeigten sich bei beiden Intensitätskombinationen ähnliche Entwicklungsverläufe der korrekten Einschätzungen (vgl. Abbildung 3.3.10, Abbildung 3.3.11 sowie Abbildung 3.3.12): Die korrekten Einschätzungen wurden zwischen dem Alter von sechs bis acht Jahren markant häufiger, blieben dann bis zum Alter von 12 Jahren auf ähnlich hohem Niveau und wurden schliesslich bei fast allen Erwachsenen beobachtet. Sowohl bei der mittleren Diskriminierungsfähigkeit als auch bei den Häufigkeiten korrekter Einschätzungen war zu beobachten, dass die Mischungsresultate von zwei gleich hellen Ursprungsflüssigkeiten deutlicher bzw. häufiger korrekt diskriminiert wurden als die Mischungsresultate der unterschiedlichen Intensitätskombination. Die genauere Analyse der Einschätzungen bei den einzelnen Mischungsresultaten ergab, dass die Unterschiede zwischen den beiden Intensitätskombinationen durch die Einschätzungen bei den physikalisch möglichen Mischungsresultaten zu Stande kamen.

Im Sinne des Ansatzes des adaptiven Denkens fragt man sich nach möglichen Einflüssen auf die Problemlösung durch die Aufgabenpräsentation. Es zeigte sich, dass das Mischungsresultat von zwei gleich hellen Flüssigkeiten, das gleich geblieben war wie die Ursprungsflüssigkeiten bzw. die Flüssigkeit im Vergleichsglas recht gut als das korrekte zu erkennen war. Das Mischungsresultat hingegen, das aus der Zugabe einer helleren Flüssigkeit zu einer dunkleren resultierte und heller war als die Ursprungsbasisflüssigkeit bzw. die Vergleichsflüssigkeit, war schwieriger zu erkennen. Da die Einschätzungen der Mischungsresultate auf der Wahrnehmungsebene bei der Mischung unterschiedlicher Intensitäten deutlich besser waren, kann dies nicht der Grund für den Unterschied sein. Woran kann es sonst gelegen haben? Wieso war der Mischungsvorgang von zwei verschieden hellen Flüssigkeiten schwieriger einzuschätzen als derjenige von zwei gleich hellen Flüssigkeiten?

Zur Beantwortung der Frage muss vermutlich berücksichtigt werden, dass dies in der Prädiktionsaufgabe ziemlich anders aussah: Bei gleichen Intensitäten konnten bei den Achtjährigen nur ca. 50 Prozent, bei den Zehnjährigen gar nur gut 30 Prozent und bei den Zwölfjährigen immerhin gut 60 Prozent das Mischungsresultat konsistent korrekt voraussagen. Die Beurteilung der Mischungsresultate der gleichen Intensitätskombination fiel den Kindern im *magic experiment* leichter als in der Prädiktionsaufgabe. Besonders eindrücklich demonstrierten dies in der Bedingung des *magic experiment* auch die inkonsistenten Regelanwender aus der Phase 1: Die zehn- und zwölfjährigen Inkonsistenten aus Phase 1 diskriminierten die beiden Mischungsresultate fast genauso deutlich wie die Durchschnittsregelanwender aus Phase 1 (Abbildung 3.3.13). Zudem urteilte die deutliche Mehrheit der vorher Inkonsistenten im *magic experiment* korrekt (Tabelle 3.3.1 und Tabelle 3.3.2).

In der Diskussion zu den Ergebnissen der Prädiktionsaufgabe war vermutet worden, dass das kognitive Schema des mentalen Ausbalancierens nicht aktiviert werden konnte und möglicherweise auch die Wahlaufgabe zur vorschnellen Anwendung des *direct functioning reasoning* geführt hatte. Beides spielte beim *magic experiment* keine Rolle mehr. Der Mischungsvorgang wurde vor Augen der Versuchspersonen vollzogen und die Einschätzungen erfolgten anhand einer 11-stufigen Ratingskala. Das bessere Abschneiden im *magic experiment* leuchtet aus mehreren Gründen ein: Der Mischungsvorgang wurde präsentiert, was vermutlich im Falle der Mischung von zwei gleich hel-

len Flüssigkeiten auf hohe Vertrautheit aus dem Alltag stiess und deshalb leicht erkannt werden konnte. Zudem waren durch die fein abgestufte Rating-Skala auch differenzierter Antworten möglich, während die Wahlaufgabe möglicherweise (vorschnell) zu falschen Urteilen geführt hatte. Wie die Auswertung der Einschätzungen auf der Wahrnehmungsebene gezeigt hatte, wurden die physikalisch korrekten Mischungsergebnisse der gleichen Intensitätskombination zwar im Vergleich mit den anderen Resultaten schlecht erkannt. Dennoch ist es möglich, dass – wenn das Mischungsergebnis als gleich geblieben taxiert wurde – es aufgrund der Tatsache, dass es *genau* gleich geblieben war, einfacher als ein realistisches Mischungsergebnis zu beurteilen war. Dies könnte auch ein Grund dafür gewesen sein, dass das Mischungsergebnis bei der unterschiedlichen Intensitätskombination, das heller geworden war, schwieriger als ein vollkommen realistisches Mischungsergebnis einzustufen war. Möglicherweise war bei den Versuchspersonen nicht die Frage zentral, ob es hätte heller werden sollen, sondern ob es im physikalisch möglichen Fall *genau* so hell hätte werden sollen, wie es geworden war. Weiter ist auch denkbar, dass das Rekapitulieren des Mischungsvorganges bei der Einschätzung des Mischungsergebnisses bei zwei gleich hellen Flüssigkeiten einfacher war als bei zwei unterschiedlichen Flüssigkeiten.

Die Entwicklungsverläufe zeigten weiter, dass in der gesamten Spanne des Grundschulalters keine eine Verbesserung der Diskriminierungsfähigkeiten stattfand. Dieses Resultat erklärt sich z. T. durch die Addierer aus Phase 1, welche die beiden Mischungsergebnisse mit zunehmendem Alter deutlicher in umgekehrter Richtung unterschieden. Zu erklären bleibt aber der ausbleibende Entwicklungsanstieg bei den Durchschnittsregelwender bei den Kindern im Alter zwischen acht und zwölf Jahren (Mittelwerte in Abbildung 3.3.13 und Abbildung 3.3.14; Häufigkeiten in Tabelle 3.3.2 und Tabelle 3.3.4). Bei der unterschiedlichen Intensitätskombination war sogar eine leichte Zunahme von indifferenten Urteilen festzustellen. Aufgrund der vorliegenden differenzierten Einzelauswertungen ist diese Frage nicht zu beantworten. Vielleicht war aufgrund der verunsichernden Aufgabe grundsätzlich nicht mehr Gewissheit bzw. Stabilität zu erwarten.

Betrachtung der Ergebnisse im Lichte der Stabilitätshypothese

Im theoretischen Kapitel zum *extinction paradigm* bzw. *magic experiment* waren zwei verschiedene Aspekte hervorgehoben worden. Der erste betrifft die Möglichkeit, durch

die Präsentation physikalisch möglicher und unmöglicher Ereignisse ein frühes Verständnis zu erfassen, das mit anderen Methoden möglicherweise nicht gefunden werden kann. Die Ergebnisse hierzu sind bereits im vorangegangenen Abschnitt diskutiert worden. Der zweite Aspekt betrifft die Stabilitätsannahme bei ausgereiftem Verständnis des zu Grunde liegenden Prinzips. Da in Studien zum Erwerb der kognitiven Durchschnittsbildung immer wieder die Additionsregel zur Lösung der Mischungsaufgabe auftauchte, interessierte zudem auch die Stabilität dieser falschen Lösungsstrategie. Hier war eine Hypothesengenerierung nicht möglich.

Es zeigte sich deutlich, dass die Durchschnittsregel viel stabiler als die Additionsregel war. Es wurde offensichtlich, dass das korrekte Wissen nach Konfrontation mit widersprüchlichen Ereignissen weniger infrage gestellt wurde als die fehlerhafte Vorstellung. Die Bedeutung dieses Ergebnisses wird im Rahmen des Kapitels 4.1.2 zur Gewissheit und Stabilität der verschiedenen Lösungsstrategien im Entwicklungsverlauf (S. 231ff.) ausführlicher diskutiert und in Zusammenhang mit ähnlichen Forschungszweigen gestellt.

Entwicklung der Stabilität der Durchschnittsregel

Gemäss der Idee des *extinction paradigm* bzw. *magic experiment* ist eine kognitive Regel bei der Darbietung widersprüchlicher Ereignisse erst dann stabil, wenn das Verständnis der Regel vollständig ausgereift ist.

Bei der Einschätzung mit der Zauberskala zeigte sich, dass sich die Stabilität der korrekten Durchschnittsregel bei den Kindern im Alter von acht bis zwölf Jahren auf einem recht hohen Stabilitätsniveau eingependelt hatte. Allerdings wird die Durchschnittsregel auch nach dem zwölften Lebensjahr noch stabiler, wie die noch höheren Werte der Erwachsenen zeigten. In der gesamten Lebensspanne des Grundschulalters war keine Veränderung bei der Einschätzung mit der Zauberskala festzustellen. Bei der detaillierten Analyse von Unterschieden zwischen den beiden Intensitätskombinationen wurde deutlich, dass die Durchschnittsregel bei der unterschiedlichen Intensitätskombination leicht instabiler war.

Erst der bewusste Umgang mit der Regel wies auf einen zusätzlichen Entwicklungsschritt in der Spanne des Grundschulalters hin: Die korrekte Regel wurde bei beiden

Intensitätskombinationen von den Durchschnittsregelanwendern aus Phase 1 mit einem Entwicklungssprung zwischen dem Alter von 8 und 10 Jahren häufiger konsistent bestätigt. Aber auch hier ist nach dem zwölften Lebensjahr noch eine weitere Stabilitätszunahme zu erwarten, da die Erwachsenen auch auf der bewussten Ebene ausschliesslich ihre korrekte Regel bestätigten. Damit kann die Hypothese aus den Ergebnissen der Phase 1 (vgl. Diskussion zu Phase 1, S. 81ff.), dass erst nach dem Alter von acht Jahren eine Durchschnittsregel entwickelt wird, die in verschiedenen Bedingungen gleichermaßen angewendet wird und auch mit entsprechender Gewissheit vertreten werden kann, gestärkt werden.

Insgesamt wurde die Durchschnittsregel im *magic experiment* sowohl bei der Diskrimination als auch – allerdings nur mit einer Tendenz – bei der Einschätzung auf der bewussten Ebene bei der Mischung gleicher Ursprungsintensitäten stabiler vertreten als bei der Mischung unterschiedlicher Intensitäten. Erklärt werden kann dieser Unterschied einerseits damit, dass die Aufgabe im *magic experiment* für die gleiche Intensitätskombination einfacher war als für die unterschiedliche Intensitätskombination. Andererseits war die Bedingung in Phase 1 für die gleiche Intensitätskombination schwieriger, so dass vermutet werden kann, dass die korrekten Einschätzungen in Phase 1 bei den gleichen Intensitäten fundierter waren als diejenigen bei den unterschiedlichen Intensitäten und sie dementsprechend auch stabiler in einer anderen Bedingung vertreten werden konnten.

Entwicklung der Stabilität der Additionsregel

In der prädiktiven Bedingung wiesen die Mischungsurteile der jüngeren Kinder bei beiden Intensitätskombinationen mit durchschnittlich gut 20 Prozent einen substanziellen Anteil an Additionsregeln auf. Der Vergleich der beiden Intensitätskombinationen ergab, dass die Additionsregeln bei der Mischungsaufgabe der gleichen Intensitätskombination etwas häufiger zur Anwendung kamen.

Hinsichtlich der Stabilität der falschen Additionsregel war eine Hypothesengenerierung nicht möglich. Grundsätzlich weiss man heute, dass viele fehlerhafte physikalische Vorstellungen erstaunlich resistent sind und z. T. bis ins Erwachsenenalter erhalten bleiben (siehe z. B. Carey, 1986; Clement, 1982; Helms & Novak, 1983; nach Sodian, 1998).

Die Fragestellung nach dem Entwicklungsverlauf im *magic experiment* war demnach mehr explorativer Art.

Wie schon früher erwähnt, kann sehr eindeutig ausgesagt werden, dass sich die Additionsregel im *magic experiment* als viel instabiler als die korrekte Durchschnittsregel erwies.

Die Entwicklungsverläufe waren hingegen recht schwierig zu interpretieren. Hinsichtlich der Diskriminierungsleistungen der Addierer aus Phase 1 zeigte sich bei der unterschiedlichen Intensitätskombination sowohl bei den Mittelwerten (**Abbildung 3.3.14**) als auch bei den Häufigkeiten (**Tabelle 3.3.4**) eine leichte Tendenz zu zunehmender Stabilität bei den älteren Kindern. Bei der gleichen Intensitätskombination konnten in der ganzen Spanne der Kinder von sechs bis zwölf Jahren keine groben Veränderungen im Urteilsverhalten ausgemacht werden.

Beim bewussten Umgang war es umgekehrt: Bei der gleichen Intensitätskombination waren die Addierer recht stabil mit einer Tendenz zu erhöhter Stabilität bei den älteren Kindern. Bei der unterschiedlichen Intensitätskombination waren die Acht- und Zwölfjährigen stabil, nicht aber die Zehnjährigen. Gründe für das abweichende Verhalten der Zehnjährigen und für die Tendenz zur Stabilisierung der fehlerhaften Intensivierungsvorstellung bei den älteren Kindern werden im Rahmen der Allgemeinen Diskussion (Kapitel 4.1.2, S. S. 231ff.) erläutert.

3.4 Kausalitätsverständnis gezauberter Ereignisse

Wie im Kapitel 2.1 zum kindlichen Verständnis gezauberter Ereignisse (S. 39ff.) dargestellt wurde, ist es möglich, nach Präsentation gezauberter und nicht-gezauberter Ereignisse Auskunft über das kindliche Kausalitätsverständnis zu erhalten. Die Ergebnisse der Untersuchung von Rosengren und Hickling (1994) liessen die Hypothese zu, dass durch die erfasste Zuschreibung der Machbarkeit zu verschiedenen Akteuren verschiedene Kausalbezüge erfasst werden können.

Es wurde deshalb im Anschluss an die Präsentation physikalisch möglicher und unmöglicher Mischungsvorgänge die Frage nach der Machbarkeit durch die verschiedenen Akteure ein anderer Zauberer, die Lehrperson, die Eltern und die Versuchsperson selbst angehängt.

3.4.1 Fragestellung

Kann mit der Machbarkeitseinschätzung das Verständnis für die Verursachung eines Ereignisses erfasst werden, so sollten die Kinder die verschiedenen Akteure insgesamt bei den physikalisch korrekten Mischungen als fähiger eingeschätzt haben als bei den unkorrekten Mischungsarten. Beim Faktor Mischungsart war entscheidend, dass die beiden Mischungsarten durch unterschiedliche Mischungsprozesse zu Stande kamen: Während die korrekten Mischungen durch simples Zusammenleeren von farbigen Flüssigkeiten (wie z. B. Sirup) entstanden waren, war für die physikalisch unmöglichen Mischungsergebnisse ein chemischer unsichtbarer Zusatz nötig gewesen. Forschungsleitend war demnach die Frage, ob sich diese Unterschiede in der Entstehungsart auch in der Zuschreibung der Machbarkeit auswirkten. Da sich die Machbarkeit bei den unkorrekten Mischungen nicht auf das blosse Zusammenleeren der Mischungen beschränkte, sondern zumindest die Beschaffung des chemischen Zusatzes implizierte, konnte man vermuten, dass diejenigen Kinder, welche die unterschiedlichen Entstehungsmechanismen erkannt hatten, auch in der Machbarkeit der verschiedenen Mischungsarten einen Unterschied machten. Allerdings war auch zu erwarten, dass die älteren Kinder zunehmend Einblick in die chemische Natur der Mischungsvorgänge hatten und deshalb auch die unkorrekten Mischungen für machbar hielten. Diese Hypothese stützte sich auf die

Ergebnisse der Studie von Rosengren und Hickling (1994), bei der sich gezeigt hatte, dass die Kinder gewöhnliche Transformationen von den aussergewöhnlichen bzw. gezauberten deutlich unterschieden: Insgesamt hielten sie die verschiedenen Akteure signifikant mehr in der Lage, die gewöhnlichen als die beiden anderen Verwandlungen zu Stande zu bringen. Die Entstehungsart des Farbmischungstricks war mit denjenigen der aussergewöhnlichen Transformationen vergleichbar.

Theoretisch besonders interessant war aber die Frage, ob sich die verschiedenen Akteure alleine und in Interaktion mit der Mischungsart bei der Machbarkeitszuschreibung unterschieden. Spiegelte sich in der eigenen Machbarkeitszuschreibung das Kausalitätsverständnis auf der Basis eigener Erfahrungen im Alltag wider, so sollten die Kinder, die korrekten Mischungen – sofern sie diese als solche erkannten – bei der eigenen Machbarkeitszuschreibung höher eingeschätzt haben als die unkorrekten Mischungsergebnisse. Die eigene Machbarkeitszuschreibung bei den korrekten Mischungsergebnissen war demnach abhängig vom Verständnis der kognitiven Durchschnittsbildung. Bei den eigenen Zuschreibungen der unkorrekten Mischungsergebnisse war zu erwarten, dass mit zunehmendem Einblick in den Entstehungsmechanismus auch das Vertrauen in sich selbst, den Trick selber reproduzieren zu können, anstieg. Die Autoritäten wie die Eltern oder die Lehrperson über ein grösseres Wissen, aber kaum über magische Fähigkeiten oder Fertigkeiten verfügen, waren bei den Zuschreibungen zu den Eltern und der Lehrperson bei beiden Mischungsarten ähnliche Entwicklungsverläufe wie bei Zuschreibungen zu sich selbst zu erwarten. Beim Zauberer hingegen war zu erwarten, dass er mit Attributen versehen wird, die ihn zur Verursachung von aussergewöhnlichen Ereignissen befähigt. Deshalb wurde angenommen, dass dem Zauberer bei beiden Mischungsarten hohe Machbarkeit zugeschrieben wurde.

Hinsichtlich des Faktors Intensitätskombination war zu erwarten, dass aufgrund der unterschiedlichen Einsicht in die kognitive Durchschnittsbildung verschiedener Intensitätskombinationen die eigenen Zuschreibungen bei diesen Faktorstufen vielleicht nicht ganz gleich ausfielen. Da aber beide Intensitätskombinationen auf dem gleichen Entstehungsmechanismus basierten, war aus dieser Sicht keine Unterscheidung zu erwarten. Interessant könnte aber der Vergleich zwischen den Farbtricks und dem Wassertrick aus dem Probeversuchsdurchgang gewesen sein. Wurde hinter den beiden Trickarten ein unterschiedlicher Mechanismus erwartet?

Auch bei den Machbarkeitszuschreibungen war es möglich, durch die Differenzwerte zwischen den Einschätzungen physikalisch möglicher und unmöglicher Mischungsresultate eine Aussage über die Diskriminationsfähigkeit zwischen den beiden Mischungsarten zu erhalten. Allerdings war dabei eine Konfundierung zwischen der Einschätzung der Physikalität der Mischungen und der Einsicht in die Entstehungsbedingungen der Mischungsresultate vorhanden. Es wurde angenommen, dass bei Einsicht in beiden Entstehungsbedingungen die beiden Mischungsarten in ihrer Machbarkeit nicht mehr unterschieden wurden. Da aber angenommen wurde, dass sich die Einsicht in die Entstehungsmechanismen der beiden Mischungsarten in der Zuschreibung zur eigenen Machbarkeit widerspiegelte, war ein Mass für die Einsicht gegeben. So sollte unter Berücksichtigung der vorhandenen Einsicht in die Entstehungsbedingungen doch noch eine Aussage über die Unterscheidungsfähigkeit der beiden Mischungsarten möglich sein. Die Unterscheidungsfähigkeit war besonders im Hinblick auf Unterschiede zwischen den verschiedenen Regelanwendern aus Phase 1 interessant. Zeigten sich auch hier Unterschiede zwischen den Durchschnitts- und Additionsregelanwendern? Zudem sollte überprüft werden, ob hier auch die beiden Intensitätskombinationen unterschieden wurden.

3.4.2 Methode

Die Angaben zu Versuchspersonen, -material und -durchführung wurden bereits in Kapitel 3.1 dargestellt. An dieser Stelle wird nur noch das Versuchsdesign mit unabhängigen und abhängigen Variablen, welches der Untersuchung des Kausalitätsverständnisses gezauberter Ereignisse zu Grunde lag, dargestellt.

Versuchsplan

Die Untersuchung des Kausalitätsverständnisses gezauberter Ereignisse bezieht sich auf Phase 3 des gesamten Experimentes, d. h. auf die Erfassung der Machbarkeitszuschreibungen zu vier verschiedenen Akteuren von physikalisch korrekten und unkorrekten Mischungen, gleicher und unterschiedlicher Ursprungsfarbindensitäten. Die folgenden Variablen lagen der Phase 3 des Experimentes zu Grunde:

Unabhängige Variablen

- Faktor A (Messwiederholungsfaktor; zweistufig): Mischungsart (physikalisch korrekt und unkorrekt)
- Faktor B (Messwiederholungsfaktor; zweistufig): Farbintensitätskombination
- Faktor C (Messwiederholungsfaktor; vierstufig): Machbarkeitszuschreibungen
- Faktor D (Zwischensubjektfaktor; fünfstufig): Alter
- Faktor E (Zwischensubjektfaktor; zweistufig): Geschlecht

Der Versuchsplan entsprach einem *2x2x4-within-subjects-design* mit vier Einschätzungen bei vier Versuchsdurchgängen pro Versuchsperson. Die Versuchsdurchgänge wurden in zwei verschiedenen Reihenfolgen dargeboten. Zu Beginn des Versuches wurde bei allen Versuchspersonen zuerst ein Probeversuchsdurchgang durchgeführt, bei welchem die Mischung von zwei farblosen Flüssigkeiten farbig (gleiches intensives Rotviolett wie bei den Experimentaldurchgängen) wurde. Die Angaben zum Alter und zur Geschlechtsverteilung der Versuchspersonen erfolgten bereits in Kapitel 3.1.

Abhängige Variablen

Machbarkeitszuschreibungen: Was meinst du, könntest du selber / könnte ... (vier Akteure) diesen Zaubertrick auch? Die Versuchsteilnehmer hatten die beiden Mischungsarten der beiden Intensitätskombinationen danach einzuschätzen, ob dieselben Resultate (nicht nur das Zusammenleeren vorhandener Flüssigkeiten) auch von einem anderen Zauberer, von ihrer Lehrperson, von ihren Eltern (als andere Autoritätspersonen) oder von ihnen selber reproduziert werden könnten. Die Antworten wurden anhand einer 5-stufigen Rating-Skala von -2 bis +2 erhoben: nein, eher nein, vielleicht, eher ja, ja. Zur statistischen Berechnung wurden die erhobenen Werte von -2 bis +2 zum Bereich von 0 bis 4 umkodiert wurden. Die mittlere Machbarkeitszuschreibung lag demnach beim Wert 2.

Erklärungen zu den Entstehungsbedingungen: Im Anschluss an die Präsentation aller Versuchsdurchgänge wurden die Kinder noch danach gefragt, ob sie eine Idee hätten, wie die Zaubertricks zu Stande gekommen waren. Die Antworten der Kinder wurden auf dem Erfassungsblatt notiert und anschliessend kategorisiert.

3.4.3 Auswertung

Gemäss Bortz und Döring, (2002, S. 223) geht man bei einer 5-stufigen Rating-Skala von -2 bis +2 davon aus, dass die Kategorien der Rating-Skala äquidistant sind. Da aber noch immer eine Kontroverse im Gange ist, ob Rating-Skalen intervall- oder ordinalskaliert sind (Bortz & Döring, 2002, S. 180f), wurden die Daten sowohl mit parametrischen als auch nonparametrischen Testverfahren ausgewertet.

Die mit einer 5-stufigen Rating-Skala erfassten Daten wurden zunächst als intervallskalierte Daten behandelt. Zu Beginn wurde eine Varianzanalyse (Allgemeines lineares Modell) für alle Machbarkeitszuschreibungen mit den Messwiederholungsfaktoren Mischungsart, Intensitätskombination und Akteure sowie den Zwischensubjektfaktoren Alter und Geschlecht gerechnet. Weiter wurde geprüft, ob die Reihenfolge der Experimentaldurchgänge (erneute Varianzanalyse mit diesem Faktor anstelle des Geschlechts) auch tatsächlich keinen Einfluss hatte.

Im Anschluss an die Varianzanalyse wurden zur Aufklärung der Haupteffekte im Anschluss an die Varianzanalyse Post-Hoc-Tests nach Scheffé gerechnet. Bei auftretenden Interaktionen zwischen verschiedenen Faktoren wurden weitere Varianzanalysen, getrennt für die Stufen eines Faktors, berechnet. Bei mehrstufigen Faktoren sollten nebst dem Post-Hoc-Test nach Scheffé noch a posteriori-T-Tests berechnet werden. Hier sollte eine Bonferoni-Korrektur vorgenommen werden.

Im Anschluss daran wurden die auf einer 5-stufigen Rating-Skala erfassten Machbarkeitszuschreibungen noch mit ordinalen Testverfahren auf Effekte von einzelnen Faktoren überprüft. Folgende Tests wurden angewendet:

- Für zwei unabhängige Stichproben: Mann-Whitney-U-Test,
- für mehrere unabhängige Stichproben: Kruskal-Wallis-H-Test,
- für zwei abhängige Stichproben: Wilcoxon-Test,
- für mehrere abhängige Stichproben: Friedman-Test.

Die Resultate dieser Berechnungen wurden aber der Übersicht halber immer in Fussnoten ausgeführt.

3.4.4 Ergebnisse

In diesem Kapitel werden zunächst die wichtigsten Ergebnisse der vierfaktoriellen Varianzanalyse in Form einer Übersicht zusammenfassend dargestellt. Direkt im Anschluss werden die einzelnen Resultate ausführlich dargestellt und eingehend ausgewertet. Danach werden die Zuschreibungen bei den physikalisch korrekten und unkorrekten Mischungen mit denjenigen beim Probeversuchsdurchgang, bei welchem aus zwei farblosen Flüssigkeiten eine farbige Mischung entstand, verglichen. In einem weiteren Teil folgt die Darstellung der Auswertungen der Diskriminierungen zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten bei den Zuschreibungen zu verschiedenen Akteuren. Schliesslich wird noch dem zwar nicht postulierten, aber inhaltlich sehr bedeutsamen Ergebnis der auftretenden Skepsis gegenüber dem Zauberer bei den Zehn- und Zwölfjährigen im Detail nachgegangen.

Die Varianzanalyse (Allgemeines lineares Modell) mit den Messwiederholungsfaktoren Mischungsart, Akteur und Intensitätskombination sowie dem Zwischensubjektfaktor Altersgruppe ergab hoch signifikante Haupteffekte bei allen Faktoren (alle Faktoren auf dem .001-Signifikanzniveau ausser dem Faktor Intensitätskombination auf dem .01-Signifikanzniveau). Die Wechselwirkung der beiden Faktoren Mischungsart und Akteur wirkte sich höchst signifikant auf die Machbarkeitszuschreibungen aus: Während der Zauberer zu beiden Mischungsarten als gleich hoch befähigt eingeschätzt wurde, zeigten sich bei den Zuschreibungen der anderen Akteure in der Reihenfolge der Versuchsperson selbst, der Eltern und der Lehrperson deutliche Unterschiede zwischen der Fähigkeit, die physikalisch korrekten und unkorrekten Mischungsarten zu reproduzieren. Zudem ergaben sich auch höchst signifikante Interaktionseffekte der Faktoren Akteur und Altersgruppe sowie Mischungsart und Altersgruppe. Weiter ergab der Effekt der Wechselwirkung der drei Faktoren Intensitätskombination, Akteur und Altersgruppe einen höchst signifikanten Wert. Die Überprüfung des Einflusses von Geschlechtszugehörigkeit sowie der Reihenfolge der Experimentaldurchgänge zeigte keine Effekte dieser Faktoren.

Bedingt durch das Auftreten der insgesamt drei Wechselwirkungen mit der Altersgruppe sowie des Haupteffektes der Intensitätskombination wurden vier weitere Varianzana-

lysen getrennt für die einzelnen Altersgruppen berechnet. Die Resultate dieser Varianzanalysen werden an dieser Stelle zusammenfassend tabellarisch dargestellt.

Tabelle 3.4.1

Zusammenfassung der Ergebnisse der vier Varianzanalysen, welche getrennt für die einzelnen Altersgruppen berechnet wurden

Altersgruppe	Haupteffekt			Interaktion
	Mischungsart	Akteur	Intensitätskombination	Akteur x Intensitätskombination
Sechsjährige	ns	***	ns	*
Achtjährige	***	***	*	ns
Zehnjährige	***	***	Trend	*
Zwölfjährige	***	***	ns	ns
Erwachsene	*	*	ns	ns

Anmerkung. Die Sternsymbole entsprechen den drei verschiedenen Signifikanzniveaus: * entspricht $p < .05$, ** entsprechen $p < .01$ und *** entspricht $p < .001$.

Diese ersten Ergebnisse werden nun in der folgenden Reihenfolge weiter ausgeführt: Machbarkeitseinschätzungen beider Mischungsarten, Machbarkeitseinschätzungen verschiedener Akteure bei allen Versuchsdurchgängen insgesamt, Machbarkeitseinschätzungen verschiedener Akteure bei verschiedenen Mischungsarten (hier folgt noch ein Exkurs zu den Machbarkeitszuschreibungen bei der Versuchsperson selbst) sowie Machbarkeitseinschätzungen bei verschiedenen Intensitätskombinationen und Altersgruppen.

Machbarkeitseinschätzungen in Abhängigkeit der beiden Mischungsarten

Die vierfaktorielle Varianzanalyse (siehe S. 173) hat einen höchst signifikanten Haupteffekt der Mischungsart ergeben, $F(1, 134) = 52.87$, $p < .001$. Zudem resultierte eine höchst signifikante Wechselwirkung der Mischungsart mit dem Alter, $F(4, 134) = 5.79$, $p < .001$. Wie aus der Abbildung 3.4.1 zu entnehmen ist, wurden die physikalisch unkorrekten Mischungsarten als weniger machbar eingeschätzt als die korrekten. Allerdings zeigten sich diese Unterschiede nicht bei allen Altersgruppen: Die Überprüfung des Effektes der Mischungsart mit der Varianzanalyse in jeder einzelnen Altersgruppe

zeigte, dass bei den Sechsjährigen kein Effekt, bei den Acht-, Zehn- und Zwölfjährigen aber ein höchst signifikanter Effekt ($F(1, 29) = 12.78$, $F(1, 29) = 20.91$, $F(1, 28) = 24.26$; jeweils mit $p < .001$) sowie bei den Erwachsenen ein signifikanter Effekt ($F(1, 19) = 4.44$, $p < .05$) vorlag.

Die Resultate der Überprüfung der Unterschiede zwischen den beiden Mischungsarten mit dem entsprechenden ordinalen Testverfahren waren alle auf dem gleichen Signifikanzniveau wie bei der Berechnung der Varianzanalyse⁶.

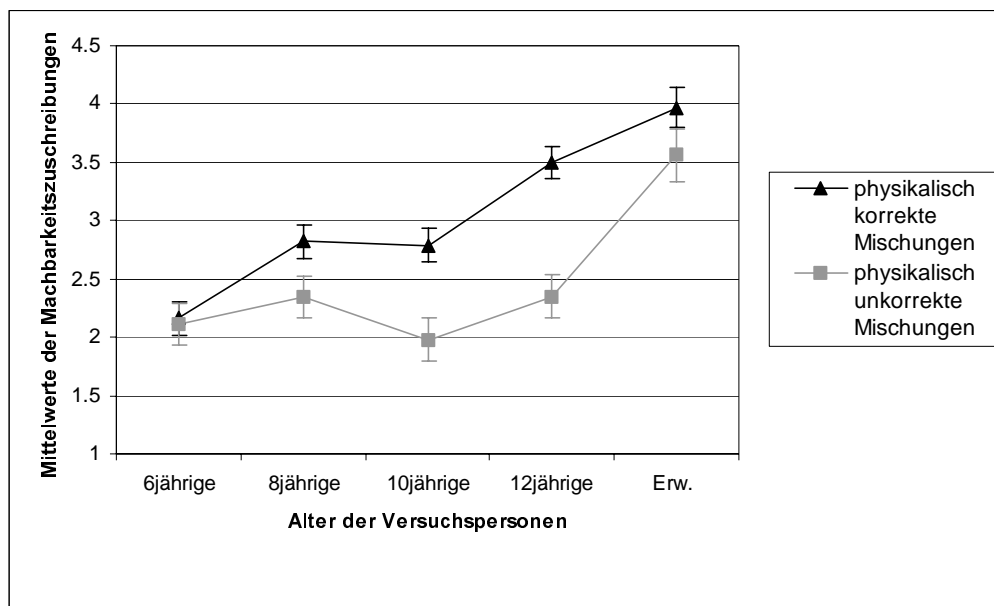


Abbildung 3.4.1. Darstellung der Mittelwerte und Standardfehler der Machbarkeitszuschreibungen aller Versuchsdurchgänge pro Altersgruppe für die beiden Mischungsarten.

Machbarkeitseinschätzungen in Abhängigkeit verschiedener Akteure bei allen Versuchsdurchgängen zusammen

Die vierfaktorielle Varianzanalyse (siehe S. 173) ergab, dass sich die verschiedenen Akteure bei den Machbarkeitszuschreibungen bei allen Versuchsdurchgängen zusammen höchst signifikant unterschieden, $F(3, 134) = 79.98$, $p < .001$. Die Mittelwerte der Zu-

⁶ Resultate der Berechnungen des Unterschieds zwischen den Zuschreibungen der beiden Mischungsarten (korrekt, unkorrekt) mit dem Wilcoxon-Test: Alle Versuchspersonen insgesamt, $z(139) = 6.51$, $p < .001$; Berechnungen pro Altersgruppe: Achtjährige: $z(30) = 3.22$, $p = .001$; Zehnjährige: $z(30) = 3.8$, $p < .001$; Zwölfjährige: $z(29) = 3.78$, $p < .001$; Erwachsene: $z(20) = 2.02$; $p < .05$.

schreibungen zeigten folgende Reihenfolge mit abnehmender Machbarkeit: Zauberer (Mittelwert: 3.69) > Lehrer (2.65) > Eltern (2.22) > Versuchsperson selbst (2.19). Zudem war auch die Interaktion mit dem Faktor Altersgruppe höchst signifikant, $F(12, 134) = 4.79, p < .001$. Die Abbildung 3.4.2 verdeutlicht die unterschiedlichen Zuschreibungen der Machbarkeit aller Versuchsdurchgänge zusammen bei verschiedenen Akteuren. Ausserdem wird auch erkennbar, dass sich die Zuschreibungen insgesamt im Altersverlauf unterscheiden.

Für die detailliertere Untersuchung des Effektes des Faktors Akteur wurden die Unterschiede in Einzelvergleichen mit t-Tests für abhängige Stichproben überprüft. Gemäss Bortz (1999) ist bei abhängigen Stichproben eine α -Fehler-Korrektur zwar nicht zwingend nötig, man kann die Resultate aber zusätzlich absichern, wenn man auch bei abhängigen Tests die α -Fehler-Adjustierung mittels Bonferoni-Korrektur einsetzt. Die drei Einzelvergleiche (Zauberer – Lehrperson, Lehrperson – Eltern, Eltern – Versuchsperson selbst) wären demnach mit $\alpha' = .05/3 = .017$ zu prüfen. Da die Ergebnisse höchst signifikante Unterschiede zwischen Zauberer und Lehrperson, $t(139) = 9.48; p < .001$ (2-seitig) sowie zwischen Lehrperson und Eltern, $t(139) = 4.96; p < .001$ (2-seitig) aufwiesen, kam die α -Fehler-Korrektur aber gar nicht zum Tragen. Der dritte Vergleich ergab kein signifikantes Resultat, d. h. bei den unterschiedlichen Werten der Akteure Eltern und Versuchsperson selbst konnte der Zufall nicht ausgeschlossen werden. Wie man in der Abbildung 3.4.2 sehen kann, waren die Zuschreibungen zu den Eltern bei den 12-jährigen und den Erwachsenen sogar tiefer als die Zuschreibungen zu den Versuchspersonen selbst. Die jüngeren Versuchspersonen hingegen trauten den Eltern die Fähigkeit, die Mischungen reproduzieren zu können, mehr zu als ihnen selbst.

Die Resultate der Berechnungen der Unterschiede der Zuschreibungen der Fähigkeiten verschiedener Akteure bei allen Versuchsdurchgängen zusammen mit dem entsprechenden ordinalen Verfahren kamen zum gleichen Ergebnis⁷.

⁷ Alle vier Akteure (Friedman-Test), $\chi^2(3, n = 139) = 174.75; p < .001$. Weitere Berechnungen mit dem Wilcoxon-Test zeigten, dass sowohl zwischen Zauberer und Lehrperson ($z(139) = 7.78; p < .001$; 2-seitig), zwischen Lehrperson und Eltern ($z(139) = 5.17; p < .001$; 2-seitig) sowie zwischen Lehrperson und Versuchsperson selbst ($z(139) = 5.2; p < .001$; 2-seitig) höchst signifikante Unterschiede bestanden.

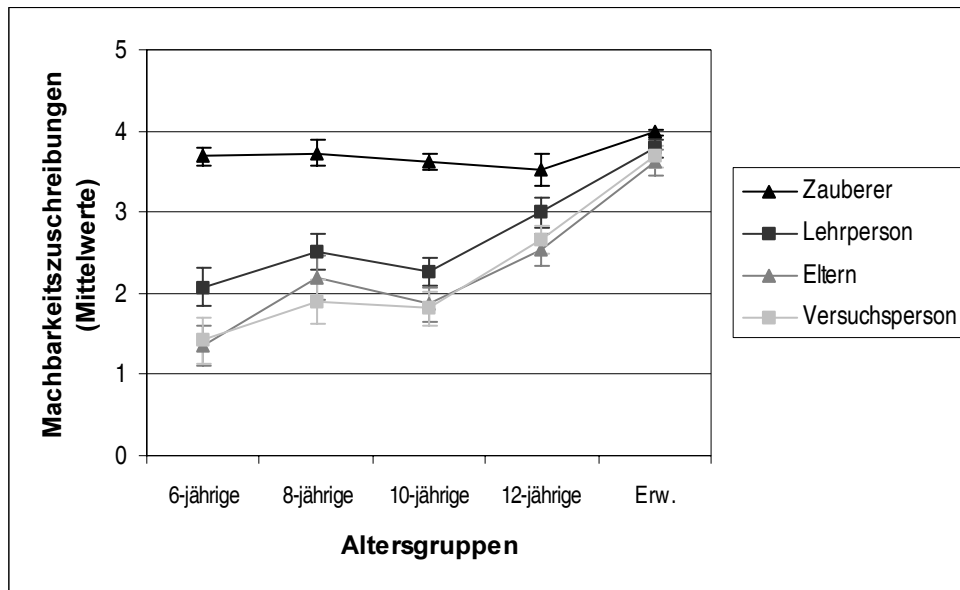


Abbildung 3.4.2. Darstellung der Mittelwerte und Standardfehler der Machbarkeitszuschreibungen aller Versuchsdurchgänge pro Altersgruppe für die verschiedenen Akteure.

Die Überprüfung der Unterschiede mit der Varianzanalyse in jeder einzelnen Altersgruppe (vgl. Tabelle 3.4.1) ergab bei allen Altersgruppen der Kinder höchst signifikante Ergebnisse (Sechsjährige: $F(3, 29) = 28.97, p < .001$; Achtjährige: $F(3, 29) = 19.33, p < .001$; Zehnjährige: $F(3, 29) = 33.12, p < .001$; Zwölfjährige: $F(3, 28) = 11.59, p < .001$. Bei den Erwachsenen unterschieden sich die Akteure zwar nicht mehr so deutlich, aber immer noch signifikant, $F(3, 19) = 3.77, p < .05$.

Erneut waren die Resultate der Überprüfung der Unterschiede zwischen den beiden Mischungsarten mit dem ordinalen Testverfahren denjenigen mit der varianzanalytischen Berechnung sehr ähnlich⁸.

Weiter wurden die Altersunterschiede der Zuschreibungen pro Akteur mit der multivariaten Varianzanalyse (ALM) überprüft. Es zeigten sich insgesamt höchst signifikante Altersunterschiede bei den vier Akteuren, $F(16, 536) = 3.55, p < .001$. Die Tests der Zwischensubjekteffekte wiesen bei allen Akteuren – ausser dem Zauberer – auf höchst signifikante Altersunterschiede in der Reihenfolge Versuchsperson selbst, Eltern und

⁸ Die Überprüfung der Unterschiede der Zuschreibungen zu verschiedenen Akteuren innerhalb der einzelnen Altersgruppe (Friedman-Test) zeigte, dass sich die vier Akteure bei allen Altersgruppen hoch signifikant unterscheiden: Sechsjährige: $\chi^2(3, n = 30) = 44.58, p < .001$; Achtjährige: $\chi^2(3, n = 30) = 38, p < .001$; Zehnjährige: $\chi^2(3, n = 30) = 59.81, p < .001$; Zwölfjährige: $\chi^2(3, n = 29) = 26.25, p < .001$; Erwachsenen: $\chi^2(3, n = 20) = 12.13, p > .01$.

Lehrperson hin: $F(4, 139) = 12.52, 11.8, 9.91$; alle mit $p < .001$. Die Überprüfung mit dem entsprechenden ordinalen Verfahren wies allerdings auch für die Altersunterschiede beim Zauberer ein signifikantes Resultat auf⁹.

Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit von der Mischungsart

Die Tabelle 3.4.2 enthält die Mittelwerte und Standardabweichungen der vierzig Gruppen gemäss dem 2 (Mischungsart) x 4 (Akteur) x 5 (Altersgruppe)-Design.

Tabelle 3.4.2

Gruppenmittelwerte und Gruppenstandardabweichungen (in Klammern) pro Akteur, Mischungsart und Altersgruppe

Altersgruppe	Akteur							
	Zauberer		Lehrperson		Eltern		Vpn selbst	
	korrekt	unkorrekt	korrekt	unkorrekt	korrekt	unkorrekt	korrekt	unkorrekt
Sechsjährige (n=30)	3.70 (.69)	3.68 (.70)	2.10 (1.39)	2.05 (1.44)	1.38 (1.50)	1.33 (1.54)	1.47 (1.61)	1.37 (1.74)
Achtjährige (n=30)	3.70 (.88)	3.77 (.82)	2.80 (1.23)	2.22 (1.34)	2.48 (1.58)	1.90 (1.64)	2.30 (1.59)	1.48 (1.66)
Zehnjährige (n=30)	3.85 (.33)	3.37 (.89)	2.73 (1.08)	1.80 (1.23)	2.30 (1.18)	1.42 (1.33)	2.28 (1.30)	1.35 (1.44)
Zwölfjährige (n=29)	3.79 (.82)	3.24 (1.43)	3.62 (.73)	2.38 (1.58)	3.19 (.97)	1.90 (1.61)	3.43 (.83)	1.90 (1.61)
Erwachsene (n=20)	4.00 (.00)	3.95 (.22)	3.95 (.22)	3.62 (.96)	3.95 (.22)	3.27 (1.39)	4.00 (.00)	3.38 (1.29)

⁹ Bei der Berechnung mit dem ordinalen Test für mehrere unabhängige Stichproben (Kruskal-Wallis-H-Test) zeigte sich, dass sich die altersabhängigen Zuschreibungen zum Zauberer zwar immer noch signifikant ($\chi^2(4, n = 139) = 10.01$; $p < .05$), aber weniger deutlich unterschieden als diejenigen zur Lehrperson, zu den Eltern und zur Versuchsperson selbst ($\chi^2(4, n = 139) = 36.52$; $p < .001$, $\chi^2(4, n = 139) = 35.76$; $p < .001$ und $\chi^2(4, n = 139) = 38.27$; $p < .001$).

Die Varianzanalyse (siehe S. 173) hat eine Wechselwirkung zwischen den Messwiederholungsfaktoren Mischungsart und Akteur ergeben, $F(3, 134) = 10.85$, $p < .001$. Bedingt durch diese Interaktion wurden, getrennt für die einzelnen Akteure, Varianzanalysen mit dem Messwiederholungsfaktor Mischungsart berechnet. Der Vergleich der physikalisch korrekten und unkorrekten Mischungen brachte bei allen Akteuren hoch signifikante Unterschiede: Versuchsperson selbst: $F(1, 138) = 39.5$, Eltern: $F(1, 138) = 33.59$, Lehrperson: $F(1, 138) = 37.49$, alle drei mit $p < .001$; Zauberer: $F(1, 138) = 9.66$, $p < .01$ ¹⁰.

Bedingt durch die Wechselwirkung wurden noch die Akteurunterschiede (Messwiederholungsfaktor), getrennt für die beiden Mischungsarten, varianzanalytisch berechnet. Es zeigten sich sowohl bei den physikalisch korrekten als auch bei den unkorrekten Mischungen höchst signifikante Unterschiede zwischen den Akteuren, $F(3, 139) = 50.41$ bzw. $F(3, 138) = 81.77$; beide mit $p < .001$. Die Reihenfolge der Akteure war aber bei den beiden Mischungsarten nicht identisch: bei den korrekten Mischungen: Eltern < Versuchsperson selbst < Lehrperson < Zauberer; bei den unkorrekten Mischungen: Versuchsperson selbst < Eltern < Lehrperson < Zauberer. Erneut führte die Berechnung mit dem ordinalen Testverfahren zu den gleichen Resultaten¹¹.

¹⁰ Die Berechnung der Unterschiede zwischen den Einschätzungen bei den beiden Mischungsarten pro Akteur und mit allen Altergruppen zusammen (Wilcoxon-Test) ergab die gleichen Ergebnisse wie bei der varianzanalytischen Berechnung, nämlich hoch signifikante Unterscheidungen bei der Lehrperson, den Eltern und der Versuchsperson selbst ($z(139) = 5.59, 5.26$ bzw. 5.69 , alle mit $p < .001$) und eine sehr signifikante Unterscheidung beim Zauberer ($z(139) = 3$, $p < .01$).

¹¹ Sowohl bei physikalisch korrekten als auch den physikalisch unkorrekten Mischungen unterschieden sich die Zuschreibungen zu den Akteuren mit dem Friedman-Test höchst signifikant ($\chi^2(3, n = 140) = 138.56$ bzw. $\chi^2(3, n = 139) = 161.37$; beide $p < .001$).

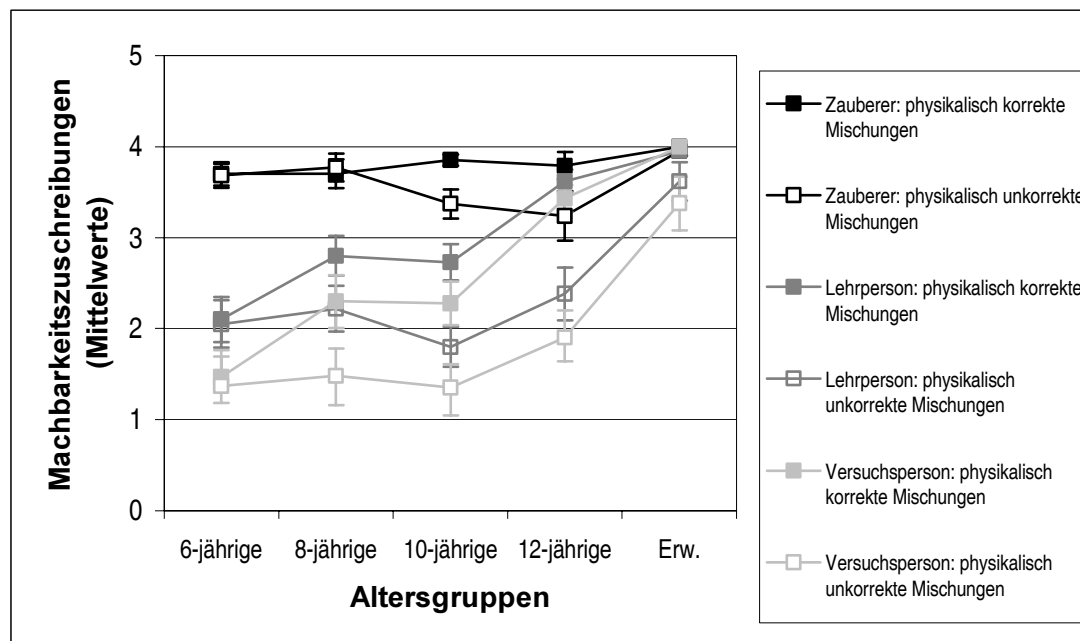


Abbildung 3.4.3. Machbarkeitszuschreibungen zu den Akteuren Zauberer, Lehrperson und Versuchsperson selbst in Abhängigkeit von der Mischungsart (korrekt, unkorrekt) und der Altersgruppe bei allen Versuchspersonen (Mittelwerte und Standardfehler).

Die grafische Darstellung der mittleren Zuschreibungen beider Mischungsarten bei den Akteuren Zauberer, Lehrperson und Versuchsperson selbst in Abbildung 3.4.3 zeigt unterschiedliche Diskriminierungen der verschiedenen Mischungen bei einzelnen Altersgruppen. Diese Unterschiede zwischen den Zuschreibungen beider Mischungsarten wurden deshalb noch, getrennt für die verschiedenen Akteure und Altersgruppen, mit Varianzanalysen auf ihre Signifikanz überprüft. Die Ergebnisse werden tabellarisch dargestellt (Tabelle 3.4.3). Obwohl mit den varianzanalytischen Ergebnissen fast identisch, werden der Vollständigkeit halber auch die Resultate mit dem entsprechenden ordinalen Testverfahren erwähnt¹².

¹² Resultate der Berechnungen des Unterschieds zwischen den Zuschreibungen bei physikalisch korrekten und unkorrekten Mischungen zu vier verschiedenen Akteuren mit dem Wilcoxon-Test für abhängige Stichproben (zweiseitige Hypothesenprüfung)

Altersgruppe	Akteur			
	Zauberer	Lehrperson	Eltern	Vpn selbst
Sechsjährige (n=30)	ns	ns	ns	ns
Achtjährige (n=30)	ns	< .01 (2.99)	< .05 (2.14)	< .01 (3.13)
Zehnjährige (n=30)	< .01 (2.86)	< .001 (3.85)	= .001 (3.39)	< .01 (3.16)
Zwölfjährige (n=29)	< .01 (2.39)	= .001 (3.46)	= .001 (3.41)	= .001 (3.42)

Tabelle 3.4.3

Resultate der Berechnungen des Unterschieds zwischen den Zuschreibungen bei physikalisch korrekten und unkorrekten Mischungen zu vier verschiedenen Akteuren pro Altersgruppe mit der Varianzanalyse

Altersgruppe	Akteur			
	Zauberer	Lehrperson	Eltern	Vpn selbst
Sechsjährige	ns	ns	ns	ns
Achtjährige	ns	< .01 (9.39)	< .05 (5.69)	= .001 (12.41)
Zehnjährige	< .01 (10.29)	< .001 (20.01)	< .001 (18.18)	= .001 (12.55)
Zwölfjährige	< .05 (6.65)	< .001 (17.80)	< .001 (18.66)	< .001 (19.83)
Erwachsene	ns	ns	< .05 (4.66)	< .05 (4.72)

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der F-Wert in Klammern dargestellt. Freiheitsgrad ist immer 1 und N bei Acht- und Zehnjährigen 29, bei Zwölfjährigen 28 und bei den Erwachsenen 19.

Aus inhaltlicher Sicht besonders erstaunlich ist die Unterscheidung der Zuschreibungen der beiden Mischungsarten beim Zauberer bei den älteren Kindern. Diese erst bei den älteren Kindern auftretende Skepsis gegenüber den Fähigkeiten, die physikalisch unkorrekte Mischung zu reproduzieren, könnte theoretisch bedeutsam sein. Deshalb wird diesem Ergebnis weiter unten (siehe Kapitel „Exkurs: Skeptische Einstellung gegenüber der Fähigkeit des Zauberers, die Tricks reproduzieren zu können“, S. 202ff.) noch ausführlicher nachgegangen.

Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit der Mischungstyp bei durchgängig korrekter Mischungsregel

Da bei obigen Berechnungen auf der Basis aller Versuchspersonen die Mischungsregeln verschieden waren, konnte angenommen werden, dass die Einschätzungen der physikalisch möglichen und unmöglichen Mischungsergebnisse nicht immer korrekt vollzogen wurden. Da aber die richtige Unterscheidung zwischen den unterschiedlichen Entstehungsbedingungen der beiden Mischungstypen eine korrekte Differenzierung von möglichen und unmöglichen Mischungen voraussetzt, war es möglich, dass die Machbar-

Erwachsene (n=20)	ns	ns	< .05 (2.04)	< .05 (2.03)
-------------------	----	----	--------------	--------------

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der Z-Wert in Klammern dargestellt.

keitseinschätzungen durch die Mischungsregel der einzelnen Versuchspersonen beeinflusst wurden¹³.

Um der Konfundierung zwischen dem Verständnis über die Entstehungsbedingungen der verschiedenen Mischungsart und dem Verständnis der kognitiven Durchschnittsbildung bei den Machbarkeitszuschreibungen zu entgehen, wurden im Anschluss an die Berechnung mit allen Versuchspersonen nur die Zuschreibungen der Anwender der Durchschnittsregel berücksichtigt. Um möglichst sicherzustellen, dass die Anwender der Durchschnittsregel tatsächlich über das Wissen der korrekten Mischungsregel verfügten, wurden nur diejenigen Versuchspersonen berücksichtigt, welche sowohl in Phase 1 als auch in Phase 2 im Sinne der Durchschnittsregel antworteten. In Abbildung 3.4.4 sind die Mittelwerte pro Altersgruppe und Akteur der Anwender der Durchschnittsregel in beiden Intensitätskombinationen abgebildet. Die Mittelwerte Altersgruppe der Sechsjährigen wurden aufgrund der kleinen Stichprobengrösse nicht abgetragen.

¹³ Ob die prädiktive Regelanwendung einen Einfluss auf die einzelnen Machbarkeitseinschätzungen hatte, wurde überprüft, indem für beide Intensitätskombinationen je drei Varianzanalysen für die drei Akteure Zauberer, Lehrperson und Versuchsperson selbst mit dem Messwiederholungsfaktor Mischungsart und den Zwischensubjektfaktoren konsistente prädiktive Regelanwendung (ohne Restkategorie) und Alter berechnet wurden. Beim Akteur Zauberer zeigte sich bei beiden Intensitätskombinationen keinen Einfluss der prädiktiven Regelanwendung. Bei der Lehrperson hingegen ergab sich bei beiden Intensitätskombinationen eine signifikante Wechselwirkung von Mischungsregel und Mischungsart (gleiche Intensitätskombination: $F(2, 122) = 3.13, p < .05$; unterschiedliche Intensitätskombination: $F(2, 125) = 4.34, p < .05$). Bei beiden war dieser Effekt dadurch zu Stande gekommen, dass die Additionsregelanwender die Mischungsarten im Vergleich mit den anderen Regelanwendern viel weniger deutlich bzw. bei der unterschiedlichen Intensitätskombination sogar leicht in umgekehrtem Sinne bei den Machbarkeitseinschätzungen unterschieden. Bei den Zuschreibungen zur Versuchsperson selbst wirkte sich die Wechselwirkung zwischen Mischungsregel und Mischungsart nur bei der gleichen Intensitätskombination signifikant auf die Einschätzungen der Machbarkeit aus ($F(2, 122) = 4.09, p < .05$).

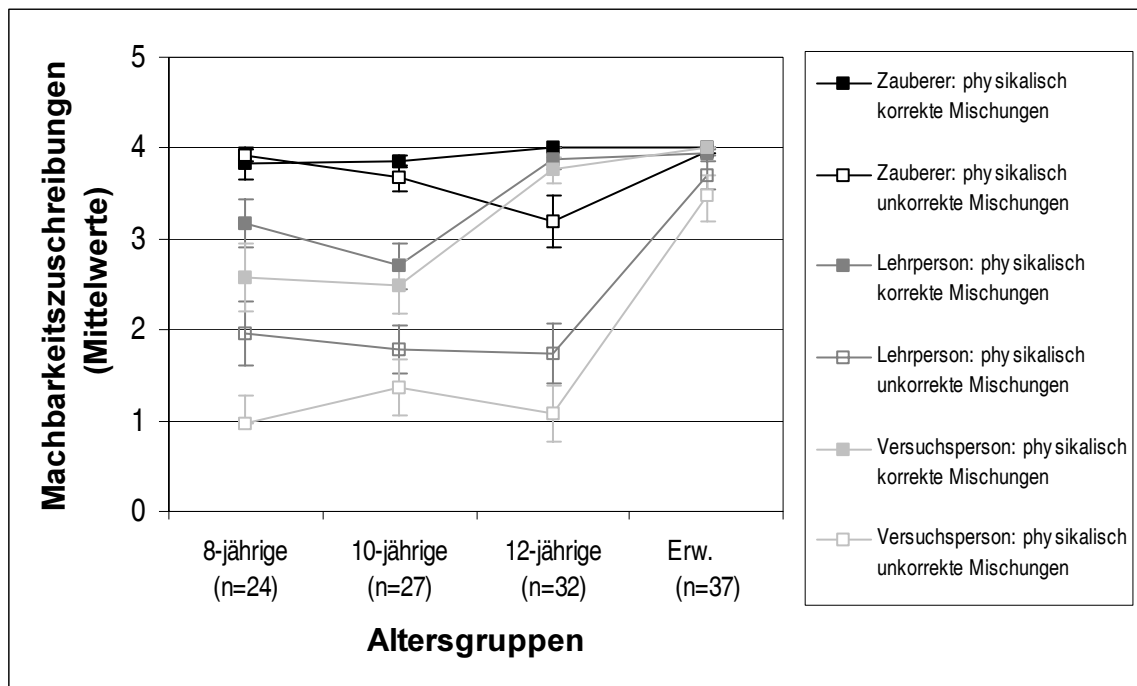


Abbildung 3.4.4. Machbarkeitszuschreibungen zu den Akteuren Zauberer, Lehrperson und Versuchsperson selbst in Abhängigkeit von der Mischungsart (korrekt, unkorrekt) und der Altersgruppe bei den Anwendern der Durchschnittsregel (Mittelwerte und Standardfehler).

Weiter wurden auf der Datenbasis der Durchschnittsregelanwender Varianzanalysen, getrennt für jede Altersgruppe und die drei Akteure Zauberer, Lehrperson und Versuchsperson selbst, mit dem Messwiederholungsfaktor Mischungsart für die Machbarkeitszuschreibungen beider Intensitätskombinationen gerechnet. Die Resultate sind in Tabelle 3.4.4 aufgetragen.

Tabelle 3.4.4

Resultate der Berechnungen des Unterschieds zwischen den Zuschreibungen bei physikalisch korrekten und unkorrekten Mischungen zu den drei verschiedenen Akteuren pro Altersgruppe mit der Varianzanalyse

Altersgruppe	Akteur					
	Zauberer		Lehrperson		Vpn selbst	
Achtjährige	ns		< .01	(8.58)	< .001	(20.93)
Zehnjährige	=.096	(2.98)	< .01	(8.87)	= .001	(13.83)
Zwölfjährige	< .01	(8.3)	< .001	(40.17)	< .001	(55.9)
Erwachsene	ns		ns		< .05	(6.85)

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der F-Wert in Klammern dargestellt. Freiheitsgrad ist immer 1 und N bei Achtjährigen 23, bei Zehnjährigen 26, bei Zwölfjährigen 31 und bei den Erwachsenen 36.

Vergleicht man die Resultate auf der Datenbasis der Durchschnittsregelanwender mit denjenigen aller Versuchspersonen, so zeigen sich keine wesentlichen Abweichungen, sondern nur geringfügige Unterschiede der Signifikanzniveaus. Vergleicht man die Darstellungen der Mittelwerte von Abbildung 3.4.3 und Abbildung 3.4.4 so zeigen sich – wie aufgrund der berechneten Konfundierung durch die Mischungsregel erwartet – leicht höhere Unterschiede zwischen den beiden Mischungsarten bei den Zuschreibungen zur Lehrperson und zur Versuchsperson selbst bei der Stichprobe der Durchschnittsregelanwender. Zusammenfassend kann aber gesagt werden, dass sich die Machbarkeitszuschreibungen durch die Konfundierung der Mischungsregel kaum verändert hatten.

Exkurs: Entwicklungsverlauf des Vertrauens in die eigene Machbarkeit der Versuchspersonen selbst

Inhaltlich kann das Vertrauen in die eigene Machbarkeit, die Mischungen selbst zu reproduzieren als Einsicht in die Entstehungsmechanismen der Mischungsarten verstanden werden. Da die Zuschreibung der Machbarkeit zur Versuchsperson selbst inhaltlich besonders bedeutsam ist, wird diese abhängige Variable an dieser Stelle noch genauer ausgewertet.

Im obigen Unterkapitel wurde bereits berichtet, dass sich die Zuschreibung zur eigenen Machbarkeit insgesamt und pro Mischungsresultat über die verschiedenen Altersgruppen hinweg bedeutsam veränderte. In Abbildung 3.4.3 ist auch zu sehen, dass die Zuschreibungen zur eigenen Machbarkeit der Zehnjährigen bei beiden Mischungsarten fast tiefer waren als diejenigen der Achtjährigen. Durch die Darstellung der Zuschreibungen anhand der Mittelwerte wurde aber verdeckt, dass sich die Zehnjährigen dennoch bei den eigenen Zuschreibungen von den Achtjährigen unterschieden. Kategorisiert man die Zuschreibungen in die drei Kategorien „minimale Zuschreibung der eigenen Machbarkeit“ (Werte 0 bis .5), „mittlere Zuschreibung der eigenen Machbarkeit“ (Werte 1 bis 3) und „maximale Zuschreibung der eigenen Machbarkeit“ (Werte 3.5 bis 4) und stellt diese grafisch dar, so lassen sich deutliche Unterschiede zwischen den Altersgruppen erkennen.

Abbildung 3.4.5 zeigt die prozentualen Anteile der drei Zuschreibungskategorien bezogen auf die Altersgruppe bei den korrekten Mischungsresultaten. Es zeigte sich, dass

sich die Zehnjährigen die eigene Machbarkeit weniger maximal zutrauten als die Achtjährigen, dafür aber vermehrt die mittlere Ausprägung der Zuschreibungen wählten¹⁴. Die minimale Zuschreibung hingegen nahm mit zunehmendem Alter stetig ab.

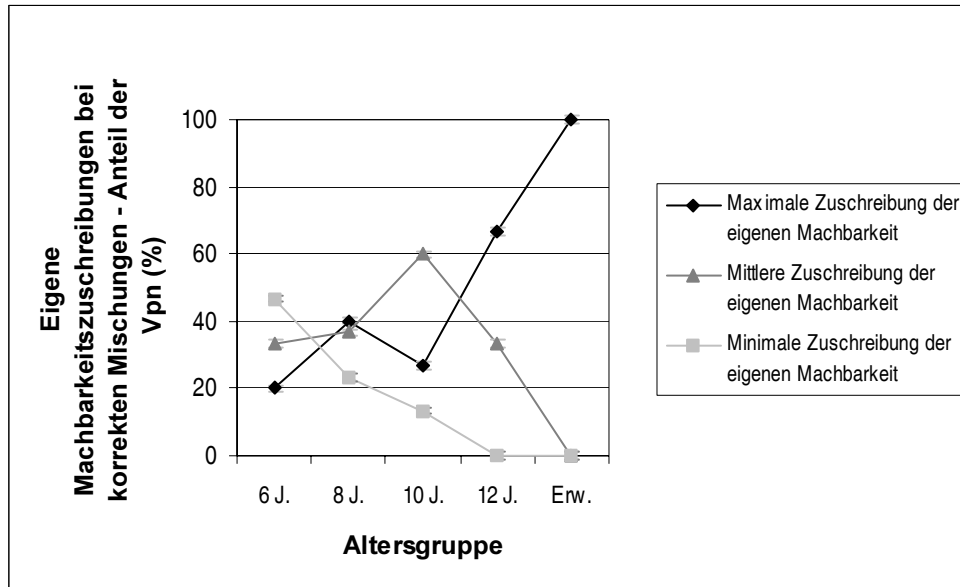


Abbildung 3.4.5. Prozentualer Anteil der Versuchspersonen der drei Kategorien von Zuschreibungen eigener Machbarkeit bei den physikalisch korrekten Mischungen auf der Datenbasis aller Versuchspersonen.

Die Verteilungen der verschiedenen Ausprägungen der eigenen Machbarkeitszuschreibungen physikalisch korrekter Mischungen über die Altersgruppen hinweg unterscheiden sich höchst signifikant, $\chi^2(8, n = 140) = 58.11; p < .001$. Statistisch besonders auffällig (d. h. höchste standardisierte Residuen bei der Berechnung der Kreuztabelle) waren die vielen minimalen und wenigen maximalen Zuschreibungen bei den Sechsjährigen, die wenigen minimalen Zuschreibungen bei den Zwölfjährigen und die ausnahmslos maximalen Zuschreibungen bei den Erwachsenen.

In Ergänzung dazu wurden zusätzlich die Verläufe auf der Datenbasis der Durchschnittsregelanwender dargestellt (Abbildung 3.4.6). Erneut zeigten sich nur geringfügig-

¹⁴ Aufgrund der erhöhten mittleren Zuschreibungen bei den Zehnjährigen wurden zusätzlich noch die Verteilungen der als „mittlere Zuschreibung“ kategorisierten drei mittleren Werte von 1 bis 3 ausgegeben. Es zeigte sich, dass alle drei Werte bei den Zehnjährigen erhöht waren, so dass keiner der drei Werte und v. a. auch nicht die mittlere Kategorie (unentschieden / weiss nicht) für die generelle Erhöhung alleine zuständig war.

gige Unterschiede zur Gesamtstichprobe: Der Unterschied zwischen mittlerer und maximaler Zuschreibungen war bei den Zwölfjährigen nicht so gross, aber auch bei den zwölfjährigen Durchschnittsregelanwendern war die mittlere Zuschreibung am häufigsten.

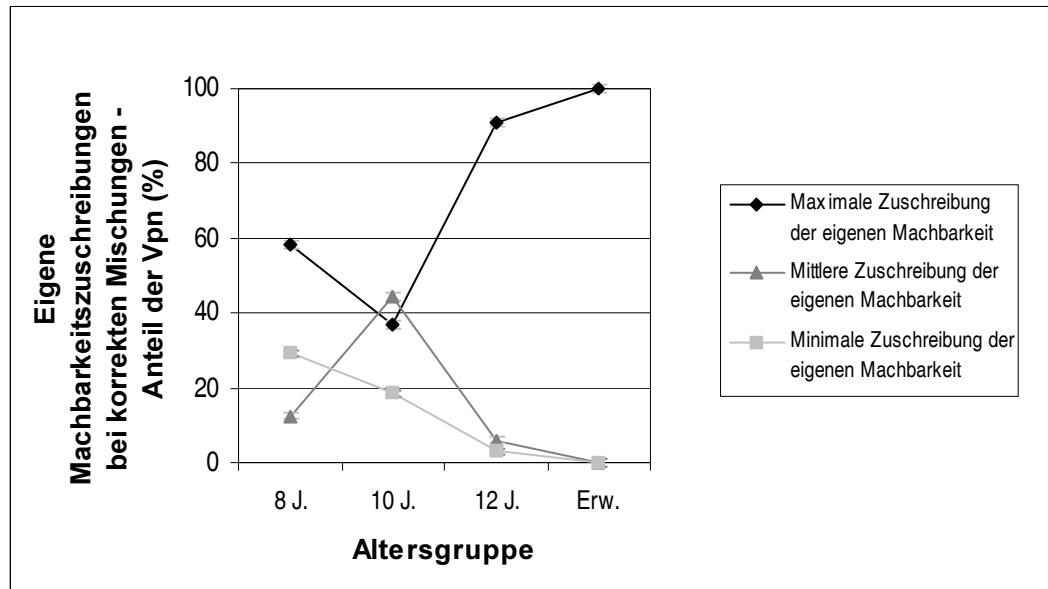


Abbildung 3.4.6. . Prozentualer Anteil der Versuchspersonen der drei Kategorien von Zuschreibungen eigener Machbarkeit bei den physikalisch korrekten Mischungen auf der Datenbasis der Durchschnittsregelanwender.

Die Überprüfung der Abhängigkeit der Verteilungen der beiden Variablen Altersgruppe und Machbarkeitszuschreibungen mit dem χ^2 -Test nach Pearson zeigte auch bei den Durchschnittsregelanwendern eine deutliche Abhängigkeit, $\chi^2(6, n = 121) = 49.64$; $p < .001$.

Abbildung 3.4.7 stellt die prozentualen Anteile der drei Zuschreibungskategorien bezogen auf die Altersgruppe bei den physikalisch unkorrekten Mischungsresultaten dar. Bei dieser Mischungsart war ebenfalls zu beobachten, dass sich die Zehnjährigen die eigene Machbarkeit weniger maximal trauten als die Achtjährigen, dafür aber vermehrt die mittlere Ausprägung der Zuschreibungen wählten¹⁵. Erneut nahm die minimale Zu-

¹⁵ Auch hier zeigte die detaillierte Analyse der Verteilungen der als „mittlere Zuschreibung“ kategorisierten drei mittleren Werte von 1 bis 3, dass alle drei Werte bei den Zehnjährigen erhöht waren, so dass keiner der drei Werte und v. a. auch nicht die mittlere Kategorie (unentschieden / weiss nicht) für die generelle Erhöhung alleine zuständig war.

schreibung hingegen mit zunehmendem Alter stetig ab, hier aber deutlich später als bei den physikalisch korrekten Mischungen.

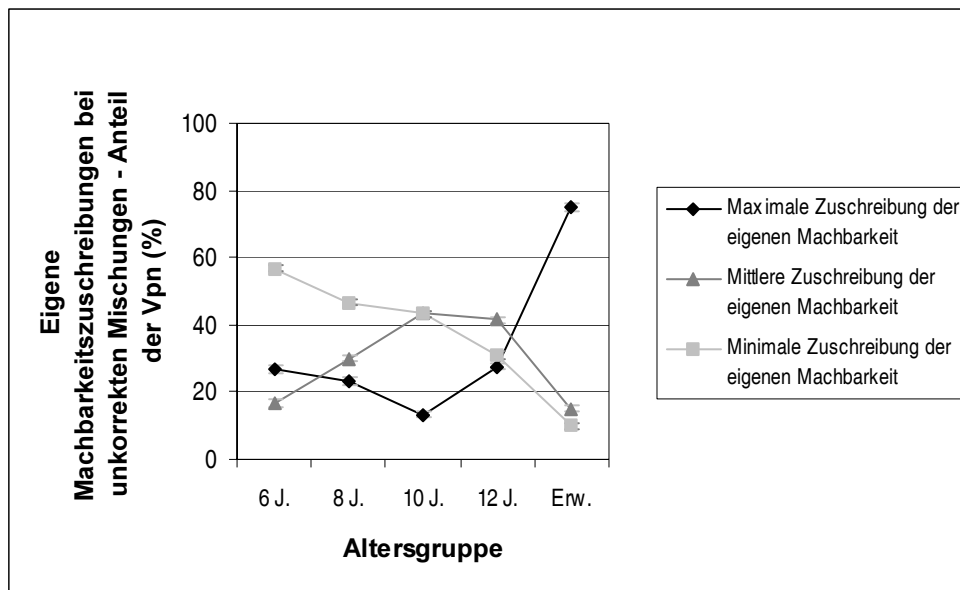


Abbildung 3.4.7. Prozentualer Anteil der Versuchspersonen der drei Kategorien von Zuschreibungen eigener Machbarkeit bei den physikalisch unkorrekten Mischungen auf der Datenbasis aller Versuchspersonen.

Auch die Verteilungen der verschiedenen Ausprägungen der eigenen Machbarkeitszuschreibungen physikalisch unkorrekter Mischungen unterschieden sich über die Altersgruppen hinweg höchst signifikant, $\chi^2(8, n = 139) = 30.69; p < .001$. Bei diesen Verteilungen waren die vielen minimalen Zuschreibungen bei den Sechsjährigen, die wenigen maximalen Zuschreibungen bei den Zehnjährigen und die wenigen minimalen dafür sehr häufigen maximalen Zuschreibungen bei den Erwachsenen statistisch besonders abweichend (höchste standardisierte Residuen bei der Berechnung der Kreuztabelle).

Auch bei der physikalisch unkorrekten Mischungsart wurden die eigenen Machbarkeitszuschreibungen noch separat für die Durchschnittsregelanwender betrachtet. In Abbildung 3.4.8 sind die Entwicklungsverläufe der kategorisierten Machbarkeitszuschreibungen dargestellt. Im Vergleich mit den Verläufen aller Versuchspersonen fällt besonders auf, dass sich die zwölfjährigen Durchschnittsbilder selbst nur gering zutrauten, die physikalisch unkorrekte Mischungsart reproduzieren zu können. Es sieht so aus, als hätten die Zwölfjährigen mit der richtigen Mischungsregel verstanden, dass sie die physikalisch korrekten Mischungen selbst reproduzieren können, nicht aber die unkor-

rekten. Die Einsicht in den Entstehungsmechanismus scheint erst bei den Erwachsenen gegeben zu sein.

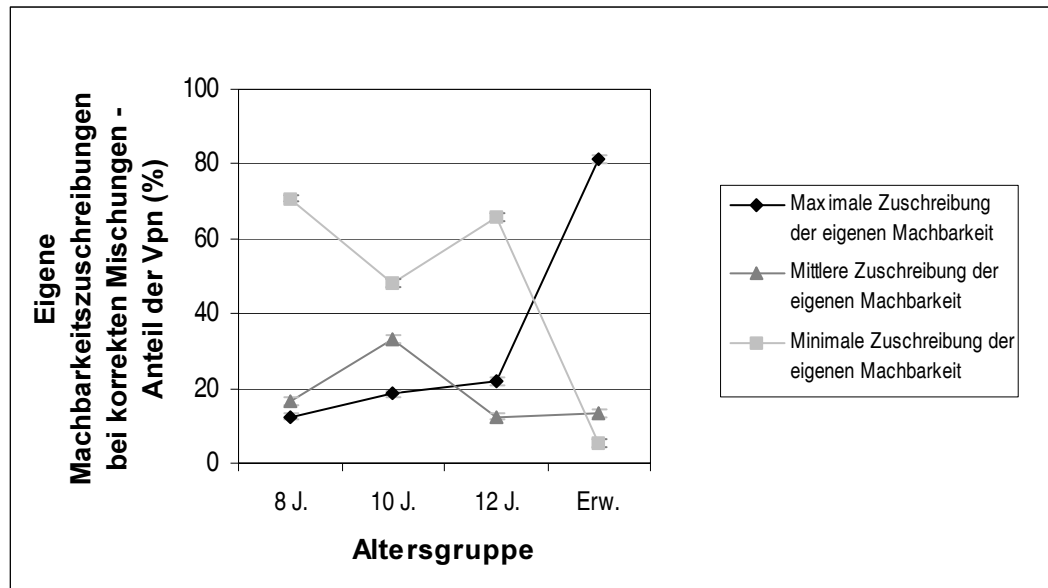


Abbildung 3.4.8. Prozentualer Anteil der Versuchspersonen der drei Kategorien von Zuschreibungen eigener Machbarkeit bei den physikalisch unkorrekten Mischungen auf der Datenbasis der Durchschnittsregelanwender.

Auch bei den physikalisch unkorrekten Mischungsergebnissen zeigte sich nach Überprüfung mit dem χ^2 -Test eine deutliche Abhängigkeit zwischen den Variablen Altersgruppe und kategorisierte Machbarkeitszuschreibungen zur Versuchsperson selbst, $\chi^2(6, n = 120) = 51.71; p < .001$.

Machbarkeitseinschätzungen in Abhängigkeit der beiden Intensitätskombinationen

Die vierfaktorielle Varianzanalyse zu Beginn der Auswertung (siehe S. 173) hatte einen sehr signifikanten Haupteffekt des Messwiederholungsfaktors Intensitätskombination ergeben, $F(1, 134) = 8.05, p < .01$. Auch dieses Resultat wurde mit dem entsprechenden Testverfahren für ordinalskalierte Daten überprüft, wobei sich dasselbe Ergebnis zeig-

te¹⁶. Abbildung 3.4.9 zeigt, dass die Mischungen gleicher Intensitätskombinationen als machbarer eingeschätzt werden als diejenigen unterschiedlicher Intensitätskombination.

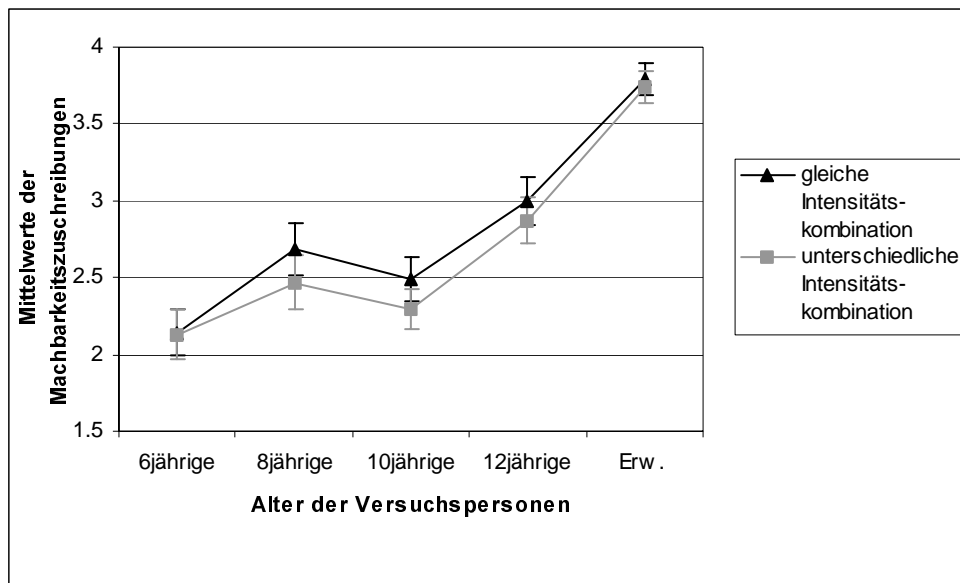


Abbildung 3.4.9. Darstellung der Mittelwerte und Standardfehler der Machbarkeitszuschreibungen aller Versuchsdurchgänge pro Altersgruppe für die beiden Intensitätskombinationen.

Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit der Intensitätskombination und der Altersgruppe

Die vierfaktorielle Varianzanalyse zu Beginn der Auswertung (siehe S. 173) ergab eine signifikante Wechselwirkung der drei Faktoren Intensitätskombination, Akteur und Alter, $F(12, 134) = 2.12, p < .05$. Bedingt durch diese Interaktion wurde die Wechselwirkung von Faktor Intensitätskombination und Faktor Akteur in getrennten Varianzanalysen für jede Altersgruppe überprüft (siehe Tabelle 3.4.1). Es zeigte sich, dass die Wechselwirkung bei Sechsjährigen und Zehnjährigen signifikant war, $F(3, 29) = 2.8, p < .05$ bzw. $F(3, 29) = 2.77, p < .05$. Die Abbildung 3.4.10 zeigt die Unterschiede der Zuschreibungen zwischen den beiden Intensitätskombinationen bei den drei Akteuren

¹⁶ Die Überprüfung des Vergleichs der beiden Intensitätskombinationen bei allen Akteuren zusammen mit dem Wilcoxon-Test zeigte einen sehr signifikanten Unterschied der Zuschreibungen, $z(139) = 2.9, p > .01$. Dabei wurden die Mischungen unterschiedlicher Intensitäten als weniger machbar eingeschätzt als diejenigen gleicher Mischungskombination. In Ergänzung dazu wurden die Intensitätsunterschiede noch getrennt für die einzelnen Altersgruppen berechnet. Es zeigte sich, dass der Unterschied zu keinem Zeitpunkt signifikant war, in den Gruppen der Acht- und Zehnjährigen aber je ein Trend zu verzeichnen war ($p = .07$ bzw. $p = .06$; Wilcoxon-Test).

Zauberer, Eltern und Versuchsperson selbst. Daraus ist zu entnehmen, dass die Sechsjährigen die Mischungen unterschiedlicher Intensitäten beim Zauberer für weniger machbar, bei den Eltern und bei sich selbst als machbarer einschätzten. Bei den Zehnjährigen zeigte sich beim Zauberer grafisch kein Unterschied bei den beiden Intensitätskombinationen, bei den Eltern und bei sich selbst hingegen wurden die Mischungen gleicher Intensitäten als deutlich machbarer beurteilt.

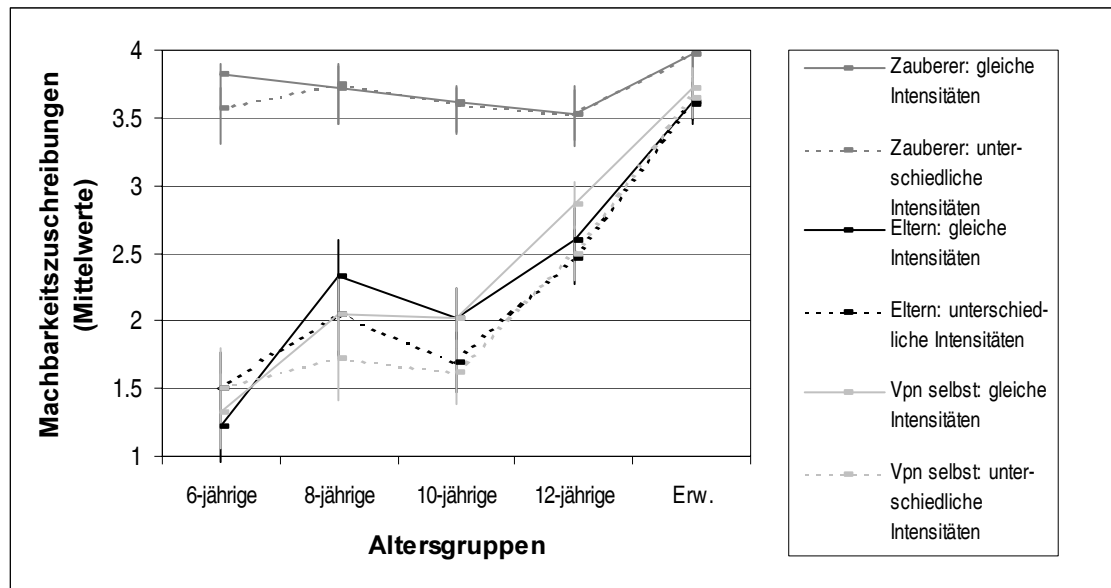


Abbildung 3.4.10. Mittelwerte und Standardfehler der Machbarkeitszuschreibungen der Akteure Zauberer, Eltern und Versuchsperson selbst in Abhängigkeit von der Altersgruppe und der Intensitätskombination.

Sowohl Abbildung 3.4.9 als auch Abbildung 3.4.10 kann entnommen werden, dass die Unterschiede zwischen den Intensitätskombinationen mit zunehmendem Alter abnehmen. Zur besseren Veranschaulichung dieses Befundes sowie der Wechselwirkung zwischen Intensitätskombination und Akteur wurden deshalb in Abbildung 3.4.11 die Differenzwerte zwischen den Zuschreibungen der gleichen Intensitätskombination und der unterschiedlichen Intensitätskombination für die vier Akteure Zauberer, Lehrperson, Eltern und Versuchsperson selbst abgetragen.

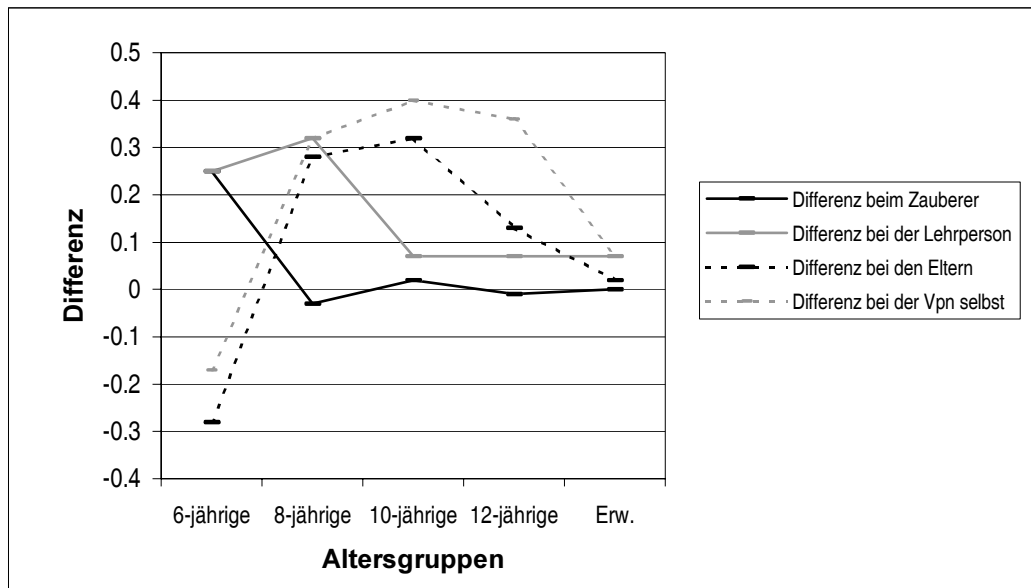


Abbildung 3.4.11. Differenzwerte zwischen den Zuschreibungen der gleichen Intensitätskombination und der unterschiedlichen Intensitätskombination für die vier Akteure (positive Werte: gleiche Intensitätskombination war machbarer; negative Werte: Mischungen unterschiedlicher Intensitäten waren machbarer).

Die obige Abbildung verdeutlicht die Interaktionen von Akteur und Intensitätskombination (dargestellt durch die Ausprägung der Differenz) bei den Sechs- und Zehnjährigen. Zudem zeigte sich ein Shift bei der Unterscheidung der Intensitätskombinationen in Abhängigkeit vom Alter und von den Akteuren: je älter die Versuchspersonen umso eher verschiebt sich das Maximum der Differenzkurve hin zu den Akteuren mit abnehmender Autorität. Diese Beobachtung wurde mit Varianzanalysen mit dem Messwiederholungsfaktor Intensitätskombination, getrennt für die Altersgruppen und die verschiedenen Akteure, statistisch überprüft.

Tabelle 3.4.5

Resultate der varianzanalytischen Berechnungen des Unterschieds zwischen den beiden Intensitätskombinationen pro Akteur und Altersgruppe

Altersgruppe	Akteur			
	Zauberer	Lehrperson	Eltern	Vpn selbst
Sechsjährige	= .05 (4.06)	ns	ns	ns
Achtjährige	ns	ns	< .05 (5.05)	ns
Zehnjährige	ns	ns	< .05 (4.42)	< .05 (6.41)

Zwölfjährige	ns	ns	ns	< .05 (5.78)
Erwachsene	ns	ns	ns	ns

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der F-Wert in Klammern dargestellt. Freiheitsgrad ist immer 1 und N bei Sechs-, Acht- und Zehnjährigen 29, bei Zwölfjährigen 28 und bei den Erwachsenen 19.

Die Ergebnisse der Varianzanalysen bestätigten die Beobachtung. Auch die Überprüfung mit dem ordinalen Testverfahren stützt diese Ergebnisse¹⁷. In der obigen Tabelle lässt sich gut ablesen, dass die signifikante Unterscheidung von oben links (d. h. jüngste Kinder und Zuschreibung beim Zauberer) nach unten rechts (d. h. älteste Kinder und Zuschreibung bei sich selbst) verlief. Die Zuschreibungen zu der Lehrperson unterschieden sich bei den beiden Intensitätskombinationen zwar nicht signifikant, in der Abbildung 3.4.11 ist aber zu erkennen, dass das Maximum der Differenzkurve bei den Achtjährigen liegt und damit die These dieses Shifts bestärkt wird.

Weitere Aufklärung lieferte der Vergleich der Intensitätsunterschiede bei den Machbarkeitszuschreibungen zwischen den physikalisch korrekten und unkorrekten Mischungsresultaten: Während bei den physikalisch korrekten Mischungsresultaten bei den drei Akteuren Lehrperson, Eltern und Versuchsperson selbst die gleiche Intensitätskombination als machbarer eingeschätzt wurde, war dies bei den unkorrekten nicht der Fall. Es scheint, als wären die physikalisch korrekten Mischungsarten, über deren Entstehungsbedingungen die Kinder zunehmend Einsicht gewinnen, bei der gleichen Intensitätskombination einfacher zu reproduzieren als diejenigen der unterschiedlichen Intensitätskombination. Bei den unkorrekten Mischungsarten wurde dieser Unterschied nur bei den Zwölfjährigen gleichermassen ausgedrückt. Obwohl auch sie kaum Einsicht in die Entstehungsbedingungen dieser Mischungsarten zeigten, könnte vielleicht doch zumindest eine Ahnung vorliegen, dass auch bei diesen Mischungen die gleichen Intensitätskombinationen einfacher zu reproduzieren sind. Bei den anderen Akteuren war bei den unkorrekten Mischungsarten zu beobachten, dass dieser Unterschied nicht gemacht wurde oder sogar die unterschiedlichen fast eher als machbarer eingeschätzt wurden. Es

¹⁷ Bei der Berechnung des Vergleichs zwischen den beiden Intensitätskombinationen mit dem Wilcoxon-Test (zweiseitige Hypothesenprüfung) resultierten die folgenden statistisch relevanten Unterschiede: Sechsjährige beim Zauberer, $z(30) = 1.89$, $p = .06$; Acht- bzw. Zehnjährige bei den Eltern, $z(30) = 2.02$, $p > .05$, bzw. $z(30) = 1.92$, $p = .05$; Zehn- bzw. Zwölfjährige bei der Versuchsperson selbst, $z(30) = 2.34$, $p > .05$, bzw. $z(29) = 2.22$, $p > .05$.

wäre möglich, dass den Eltern und den Lehrpersonen, denen die unkorrekten Mischungsarten auch nur leicht mehr zugetraut werden als sich selbst, doch eher in der Lage sein könnten, auch die Mischungen unterschiedlicher Intensitätskombinationen zu reproduzieren.

Machbarkeitseinschätzungen beim Probeversuchsdurchgang „Wassermischungstrick“ und Vergleich mit den beiden Mischungsarten des Experimentes

Bisherige Auswertungen haben gezeigt, dass sich bei beiden Mischungsarten Alters- und Akteurunterschiede bei den Machbarkeitszuschreibungen ergeben haben. Da beim Probeversuchsdurchgang noch eine dritte Variante von Mischungsart präsentiert wurde, dem so genannten „Wassermischungstrick“, bei welchem aus der Mischung von zwei farblosen Flüssigkeiten eine farbige Mischung resultierte, drängt sich ein Vergleich mit dem Wassermischungstrick auf.

Zunächst wurde eine Varianzanalyse (ALM) für die Machbarkeitszuschreibungen beim Wassertrick mit dem Messwiederholungsfaktor Akteur und dem Zwischensubjektfaktor Alter berechnet. Diese ergab höchst signifikante Haupteffekte bei den beiden Faktoren Akteur ($F(1, 133) = 64.34, p > .001$) und Alter ($F(4, 133) = 11.08, p > .001$) sowie eine sehr signifikante Wechselwirkung zwischen diesen beiden Faktoren ($F(12, 133) = 2.37, p > .01$). Der Post-Hoc-Tests nach Scheffé mit dem Vergleich zwischen den Altersgruppen zeigte, dass sich alle Altersgruppen im Vergleich mit den Erwachsenen höchst signifikant unterschieden. Es zeigte sich also ein sehr ähnliches Bild wie beim Farbmischungstrick (Siehe S. 179f.): Die Kinder schätzten den Zauberer als deutlich fähiger ein als die anderen Akteure, was bei den Erwachsenen dann nicht mehr der Fall war. Zudem stiegen die Machbarkeitszuschreibungen generell mit zunehmendem Alter an.

Interessanter ist nun aber der Vergleich mit den beiden Mischungsarten des Experimentes, den physikalisch korrekten und unkorrekten Farbmischungen. Unterschieden sich die Zuschreibungen beim deutlichen Wassermischungstrick von denjenigen Mischungen, welche nicht getrickst waren? Unterschieden sich die Machbarkeitszuschreibungen beim Wassermischungstrick von denjenigen beim Farbmischungstrick (physikalisch unkorrekte Mischungsart)? Letzterer Fragestellung wird an dieser Stelle zuerst nachgegangen.

Zwei verschiedene Tricks (Wasser- und Farbmischungstrick)

Die Berechnung der Varianzanalyse (ALM)¹⁸ mit dem Messwiederholungsfaktor Trickart (Wasser- und Farbmischungstrick) und Akteur sowie dem Zwischensubjektfaktor Alter ergab höchst signifikante Haupteffekte von Akteur ($F(3, 132) = 93.65, p < .001$) und Alter ($F(4, 132) = 11.54, p < .001$). Bei den Wechselwirkungen waren sowohl diejenige von Akteur und Alter ($F(12, 132) = 3.12, p < .001$) als auch diejenige von Trickart und Alter ($F(4, 132) = 2.46, p < .05$) signifikant. Bei der Darstellung der Mittelwerte der Machbarkeitszuschreibungen in Abbildung 3.4.12 wurden die drei Akteure Lehrperson, Eltern und Versuchsperson selbst zusammengefasst, da deren Werte sich nicht signifikant unterschieden. Das Liniengramm verdeutlicht die varianzanalytischen Ergebnisse: Der Zauberer unterschied sich von den anderen Akteuren in der Machbarkeit (HE Akteur) und die übrigen Akteure wurden mit zunehmendem Alter als fähiger eingeschätzt (HE Alter). Allerdings wurde die Differenz zwischen dem Zauberer und den übrigen Akteuren mit ansteigendem Alter kleiner (IA Akteur und Alter). Die Varianzanalysen mit den Messwiederholungsfaktoren Akteur und Trickart ergaben bei allen Gruppen der Kinder höchst signifikante Haupteffekte durch die verschiedenen Akteure, während dieser bei den Erwachsenen noch auf dem .05-Level signifikant war. Die Wechselwirkung zwischen der Trickart und dem Alter ist darauf zurückzuführen, dass nur die Achtjährigen einen Unterschied zwischen den beiden Tricks machten. Die Varianzanalysen mit den Messwiederholungsfaktoren Akteur und Trickart zeigten, dass sich die Trickart nur bei den Achtjährigen auf das Machbarkeitsurteil verschiedener Akteure auswirkte, $F(1, 29) = 5.04, p < .05$.

¹⁸ Auf eine vergleichende Berechnung mit ordinalen Testverfahren wurde in diesem Kapitel verzichtet, da sich bei den oben ausgeführten Vergleichen fast identische Resultate durch die beiden Verfahren gezeigt hatten. Dadurch war der Aufwand für eine zusätzliche Überprüfung nicht gerechtfertigt und zudem waren die Ergebnisse übersichtlicher darstellbar.

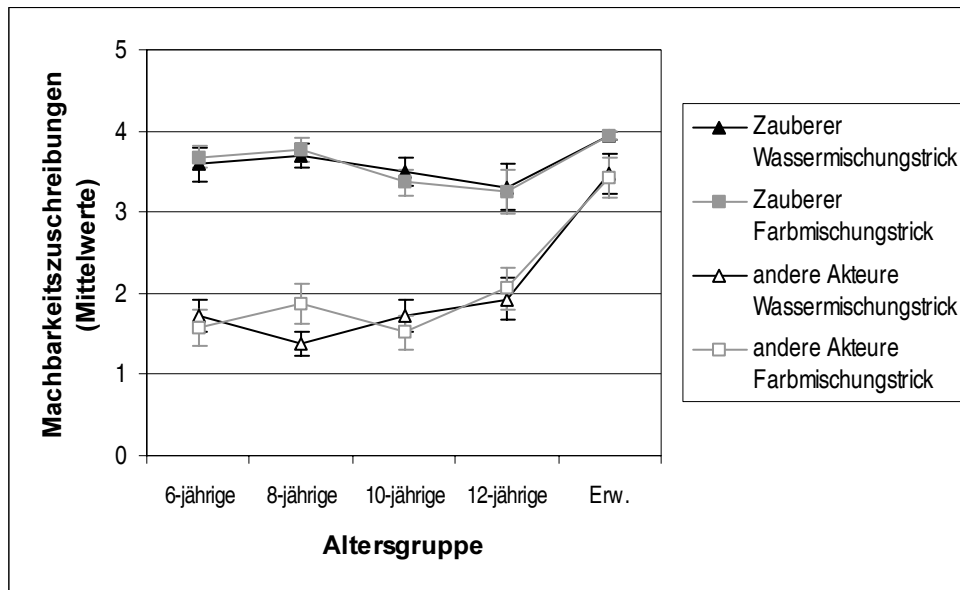


Abbildung 3.4.12 Darstellung der Mittelwerte und Standardfehler der Machbarkeitszuschreibungen bei dem Wasser- und Farbmischungstrick pro Altersgruppe für die Akteure Lehrperson, Eltern und Versuchsperson zusammen (andere Akteure) sowie den Zauberer alleine.

Aufgrund des Effektes der Trickart bei den Achtjährigen sowie dem besonderen Interesse an Zuschreibungsunterschieden beim Zauberer, wurden noch zusätzlich Varianzanalysen mit dem Faktor Trickart, getrennt für den Akteur Zauberer und die übrigen drei Akteure zusammen sowie getrennt für jede Altersgruppe, durchgeführt. Die Ergebnisse sind in der linken Hälfte der Tabelle 3.4.6 dargestellt. Es zeigte sich, dass die sich bei den Achtjährigen die beiden Tricks nur bei den Zuschreibungen der drei anderen Akteure unterschieden. Der Abbildung 3.4.12 kann man entnehmen, dass dabei der Wassermischungstrick als weniger machbar als der andere Farbmischungstrick eingeschätzt wurde. Abgesehen davon, unterschieden sich die verschiedenen Tricks der Zuschreibung zum Zauberer und zu den übrigen drei Akteuren nicht. Die Berechnungen mit dem entsprechenden ordinalen Testverfahren führten erneut zum gleichen Resultat¹⁹.

¹⁹ Die ordinale Überprüfung der Unterschiede durch die beiden Tricks bei den Machbarkeitseinschätzungen des Zauberers sowie den drei übrigen Akteuren zusammen ergab mit dem Wilcoxon-Test einzig bei den Achtjährigen und den Zuschreibungen bei den drei Akteuren ein signifikantes Resultat, $z(30) = 2.19$, $p < .05$. Zudem wurden innerhalb der Gruppe der Achtjährigen noch Vergleiche der Urteile beim Wassermischungstrick und den beiden einzelnen Versuchsdurchgängen (gleiche und unterschiedliche Intensitätskombination) durchgeführt. Dabei zeigte sich, dass die Diskrepanz durch das Mischungsresultat bei den gleichen Intensitäten zustande kam, $z(30) = 2.18$, $p < .05$. Der Wassermischungstrick sowie der Farbmischungstrick bei der unterschiedlichen Intensitätskombination unterschieden sich nicht signifikant.

Tabelle 3.4.6

Resultate der Berechnungen der Unterschiede zwischen dem Wasser- und Farbmischungstrick sowie dem Wassermischungstrick und der korrekten Mischung für den Akteur Zauberer und die übrigen Akteure zusammen mit der Varianzanalyse

Altersgruppe	Wasser- vs. Farbmischungstrick		Wassermischungstrick vs. korrekte Mischung		
	Zauberer	Übrige Akteure	Zauberer	Übrige Akteure	
Sechsjährige	ns	ns	ns	ns	
Achtjährige	ns	< .05 (5.24)	ns	< .001	(3.77)
Zehnjährige	ns	ns	< .05 (2.05)	< .01	(2.72)
Zwölfjährige	ns	ns	ns	< .001	(4.27)
Erwachsene	ns	ns	ns	= .07	(1.84)
Alle Vpn	ns	ns	< .05 (2.37)	< .001	(6.18)

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der F-Wert in Klammern dargestellt. Freiheitsgrad ist bei allen Resultaten 1 und n bei den Sechs- 28, bei den Acht- und Zehnjährigen 29, bei den Zwölfjährigen, bei den Erwachsenen 19 und bei allen Vpn 137.

Wassermischungstrick und physikalisch korrekte Mischungen

Die Berechnung der Varianzanalyse (ALM) mit dem Messwiederholungsfaktoren Realitätscharakter der Mischungen (Wassermischungstrick und physikalisch korrekte Mischung) und Akteur sowie dem Zwischensubjektfaktor Alter ergab höchst signifikante Haupteffekte aller Faktoren (Realitätscharakter der Mischungen: $F(1, 133) = 52.86$; Akteur: $F(3, 133) = 79.71$, Alter: $F(4, 133) = 19.42$; $ps < .001$). Die Darstellung der Mittelwerte in Abbildung 3.4.13 verdeutlicht die Bedeutung dieser Haupteffekte: die Machbarkeitszuschreibungen stiegen mit zunehmendem Alter an und die Akteure unterschieden sich bei den Kindern deutlich.

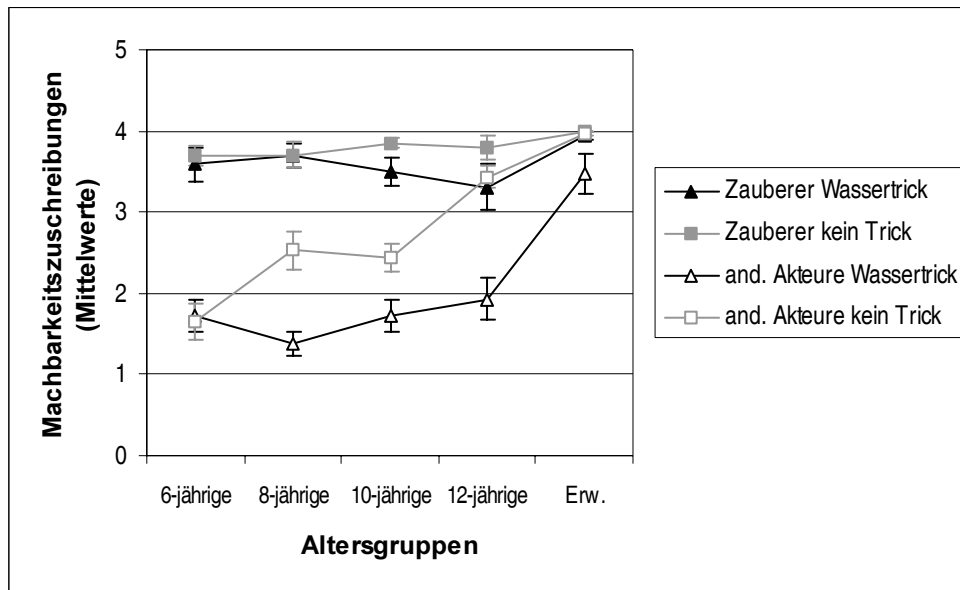


Abbildung 3.4.13. Darstellung der Mittelwerte der Machbarkeitszuschreibungen bei dem Wassermischungstrick und den physikalisch korrekten Mischungen pro Altersgruppe für die Akteure Lehrperson, Eltern und Versuchsperson zusammen (andere Akteure) sowie den Zauberer alleine.

Erkennbar wird auch, dass die Wechselwirkungen der Faktoren bedeutsam waren. Die Varianzanalysen ergaben denn auch drei höchst signifikante Interaktionen (alle mit $p < .001$): Realitätscharakter der Mischungen und Alter ($F(4, 133) = 6.51$), Akteur und Alter ($F(12, 133) = 4.25$) sowie Realitätscharakter und Akteur ($F(3, 133) = 8.73$). Zudem ergab die Wechselwirkung aller drei Faktoren einen signifikanten Effekt ($F(12, 133) = 1.97$, $p < .05$). Für das Verständnis der Bedeutungen dieser Wechselwirkungen kann erneut Abbildung 3.4.13 zur Hilfe genommen werden: Eine Wechselwirkung von Realitätscharakter der Mischungen und Alter ergab sich, da die Sechsjährigen und die Erwachsenen im Gegensatz zu den älteren Kindern die beiden Mischungsorten kaum unterschieden; wie schon beim Vergleich zwischen den beiden Mischungsorten (siehe S. 194) war die Interaktion von Akteur und Alter dadurch zu Stande gekommen, weil sich die Machbarkeitszuschreibungen beim Zauberer von denjenigen der anderen Akteure bei den Kindern, nicht aber bei den Erwachsenen unterschieden; dass der unterschiedliche Realitätscharakter der beiden Mischungen bei den Zuschreibungen zu verschiedenen Akteuren nicht als nicht immer gleich gross eingeschätzt wurde, z. B. daran zu erkennen, dass die Unterschiede beim Zauberer kleiner sind als diejenigen bei den übrigen Akteuren. Die dreifache Wechselwirkung entstand dadurch, dass diese Unterschiede noch zusätzlich bei verschiedenen Altersgruppen variierten.

Bedingt durch die mehrfachen Wechselwirkungen mit dem Faktor Alter wurden für jede Altersgruppe getrennte Varianzanalysen gerechnet und damit die obigen Beschreibungen der Abbildung 3.4.13 statistisch bestätigt. Die Resultate sind in Tabelle 3.4.7 dargestellt.

Tabelle 3.4.7.

Darstellung der Ergebnisse der Varianzanalysen (ALM) mit den Messwiederholungsfaktoren Realitätscharakter der Mischung und Akteur. Die Berechnungen wurden für jede einzelne Altersgruppe getrennt vorgenommen

Altersgruppe	Wassermischungstrick vs. korrekte Mischung					
	Realitätscharakter		Akteur		Interaktion Akteur x Realitätscharakter	
Sechsjährige	ns		< .001 (22.41)		ns	
Achtjährige	< .001	(26.68)	< .001	(24.43)	< .001	(7.54)
Zehnjährige	< .01	(8.81)	< .001	(34.83)	ns	
Zwölfjährige	< .001	(32.66)	< .001	(12.04)	< .01	(4.46)
Erwachsene	= .06	(3.83)	< .05	(3.20)	< .05	(3.18)

Anmerkung. In der linken Spalte ist jeweils der p-Wert und in der rechten der F-Wert in Klammern dargestellt. Freiheitsgrade sind beim Faktor Realitätscharakter 1, beim Akteur 3 und bei der Interaktion von Akteur und Realitätscharakter 3. N ist bei den Sechs- und Zwölfjährigen 28, bei den Acht- und Zehnjährigen 29 und bei den Erwachsenen 19.

Ausserdem wurde mit der Varianzanalyse, getrennt für die einzelnen Altersgruppen, noch detaillierter untersucht, ob sich beim Akteur Zauberer Unterschiede bei den Mischungen mit verschiedenem Realitätscharakter zeigte. Wie schon beim Vergleich der physikalisch korrekten mit den unkorrekten (siehe Tabelle 3.4.3) war auch hier ein Unterschied bei den Zehnjährigen – und nur bei diesen – zu beobachten, $F(1, 29) = 4.83$, $p < .05$). Wurden die Altersgruppen der Zehn- und Zwölfjährigen zusammen varianzanalytisch hinsichtlich des Unterschieds bei der Machbarkeitszuschreibungen beim Zauberer bei den Mischungen mit unterschiedlichem Realitätscharakter untersucht, so ergab sich ebenfalls ein signifikanter Unterschied, $F(1, 58) = 5.93$, $p < .05$. Die Unterschiede

bei den Machbarkeitszuschreibungen zwischen den verschiedenen Mischungen wurden zusätzlich noch mit dem ordinalen Testverfahren berechnet²⁰.

Ausmass der Diskriminierung zwischen den beiden Mischungsarten bei den eigenen Machbarkeitszuschreibungen

Im Abschnitt zu den Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit der Mischungsart (S. 178ff.) wurde bereits deutlich, dass sich die Unterscheidungen zwischen den beiden Mischungsarten bei den Kindern im Altersverlauf systematisch veränderten: Die Unterschiede und damit die Diskriminationen wurden mit zunehmendem Alter der Kinder grösser. Da die Erwachsenen im Gegensatz zu den Kindern kaum mehr diskriminierten, besteht der Normwert bei den Machbarkeitsdiskriminierungen nicht in einer möglichst maximalen Ausprägung. Bei Einsicht in den Entstehungsmechanismus der beiden Mischungsarten, welche bei den Erwachsenen mehrheitlich vorausgesetzt werden kann (vgl. das hohe Vertrauen in die eigene Machbarkeit der physikalisch unmöglichen Mischungen bei den Erwachsenen, S. 184ff.), werden die beiden Mischungsarten als gleich machbar eingeschätzt. Die Ausprägungen der Diskriminationen bei den Machbarkeitszuschreibungen können demnach nicht direkt mit denjenigen bei der Zauberskala verglichen werden, da bei der Machbarkeitsunterscheidung der beiden Mischungsarten eine Konfundierung mit der Einsicht in den Entstehungsmechanismus bestand. Da aber aus Abbildung 3.4.4 (siehe S. 183) deutlich hervorgeht, dass die postulierte Einsicht in die Entstehungsbedingungen der beiden Mischungsarten bei der eigenen Zuschreibungen der Kinder aus verschiedenen Altersgruppen ziemlich stabil war, können die Diskriminationen mit der Machbarkeitsskala dennoch interpretiert werden. Die gleiche Abbildung macht auch deutlich, dass bei den Erwachsenen durch die hohe eigene Machbarkeitszuschreibung zur physikalisch unkorrekten Mischungsart tatsächlich eine Konfundierung vorlag.

²⁰ Die Berechnung mit dem Wilcoxon-Test kam beim Zauberer zum gleichen Ergebnis, einem Unterschied bei den Zehnjährigen, $z(30) = 2.05$, $p < .05$. Die Berechnung mit den beiden Altersgruppen der Zehn- und Zwölfjährigen zusammen ergab erneut einen signifikanten Effekt des Realitätscharakter der Mischungen, $z(59) = 2.42$, $p < .05$. Ferner wurden noch die Unterschiede bei den übrigen drei Akteuren statistisch überprüft (Wilcoxon-Test): Sechsjährige ns; Achtjährige, $z(30) = 3.77$, $p < .001$, Zehnjährige, $z(30) = 2.72$, $p < .01$, Zwölfjährige, $z(29) = 4.27$, $p < .001$, Erwachsene, $z(20) = 1.84$, $p = .07$.

Für die Berechnungen der Diskriminationsausprägungen zwischen den beiden Mischungsarten wurden neue Variablen aus den Differenzen der eigenen Machbarkeitszuschreibungen bei den physikalisch korrekten und unkorrekten Mischungen für das 2 (Intensitätskombinationen) x 4 (Akteure)-Design gebildet. Positive Diskriminationswerte besagen eine höhere Machbarkeitszuschreibung bei den physikalisch korrekten Mischungen und negative Werte höhere Machbarkeitszuschreibungen bei den physikalisch unkorrekten Mischungen.

Mit diesen neuen Diskriminationsvariablen wurde eine Varianzanalyse (Allgemeines lineares Modell) mit dem Messwiederholungsfaktor Intensitätskombination sowie dem Zwischensubjektfaktor Alter gerechnet. Es resultierte ein Haupteffekt des Alters ($F(4, 134) = 3.62, p < .01$), aber kein Effekt der Intensitätskombination. Der Post-Hoc-Test nach Scheffé wies auf einen signifikanten Unterschied zwischen den Sechs- und Zwölfjährigen hin (mittlere Differenz = 1.42, $p = .01$).

Diskriminationen in Abhängigkeit von der prädiktiven Regel und vom Alter

Für die folgende Berechnung der Abhängigkeit der prädiktiven Regel und des Alters bei der Machbarkeitsdiskrimination wurden die Zuschreibungen zur Versuchsperson selbst ausgewählt. Da sich in Abbildung 3.4.4 (siehe S. 183) gezeigt hatte, dass Einsicht in die Entstehungsbedingungen der beiden Mischungsarten aufgrund der eigenen Machbarkeitszuschreibungen bei den Kindern ziemlich stabil war, konnte der Effekt dieser konfundierenden Variablen als einigermaßen konstant angenommen werden. Im Anschluss an die Varianzanalysen mit den Zwischensubjektfaktoren Alter und konsistente Regelanwendung (ohne Restkategorie) wurden, falls sich signifikante Wechselwirkungen ergaben, jeweils weitere Varianzanalysen mit den einzelnen Regelkategorien und den einzelnen Altersgruppen gerechnet.

Die Abhängigkeit der Machbarkeitszuschreibungen vom Alter und der prädiktiven Regelanwendung wurde separat für die einzelnen Intensitätskombinationen überprüft.

Bei der gleichen Intensitätskombination zeigten sich signifikante Haupteffekte der Faktoren Alter ($F(4, 134) = 2.46, p < .05$) und prädiktive konsistente Regelanwendung ($F(2, 134) = 4.09, p < .05$). Der Post-Hoc-Test nach Scheffé ergab bei den Unterschieden zwischen den Mittelwerten der Durchschnitts- und Additionsregelanwender eine Signi-

fikanz (mittlere Differenz = 1.06; $p < .05$) und bei den inkonsistenten Anwendern und den Addierern ein Trend (mittlere Differenz = 1.01; $p = .088$). Abbildung 3.4.14 veranschaulicht die Ergebnisse bei der gleichen Intensitätskombination grafisch.

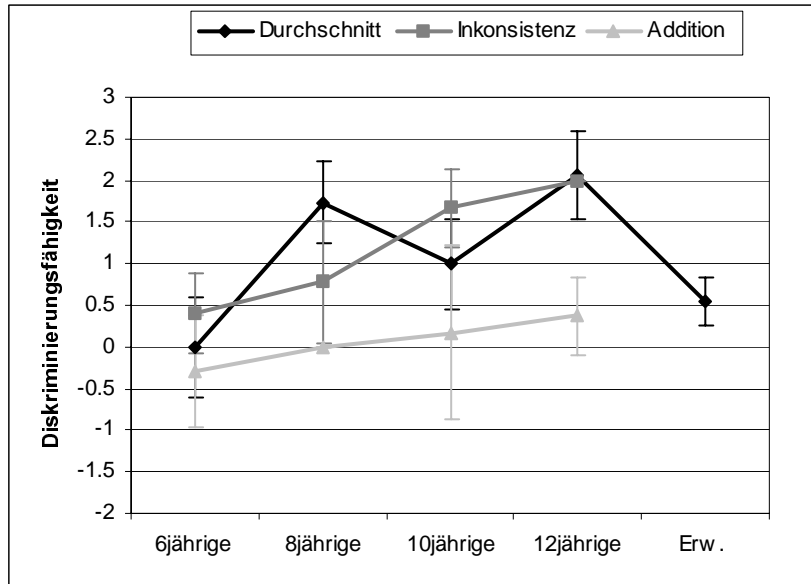


Abbildung 3.4.14. Mittelwerte und Standardfehler der Diskriminationen der eigenen Machbarkeitszuschreibungen bei der gleichen Intensitätskombination in Abhängigkeit vom Alter und der konsistenten prädiktiven Regelanwendung. Positive Diskriminationswerte besagen eine höhere Machbarkeitszuschreibung bei den physikalisch korrekten Mischungen und negative Werte höhere Machbarkeitszuschreibungen bei den physikalisch unkorrekten Mischungen.

Die Verläufe bei den Altersgruppen der Kinder zeigen bei allen Regelanwendern eine höhere Unterscheidung zwischen den beiden Mischungsarten mit zunehmendem Alter. Allerdings diskriminieren die Addierer stets tiefer, d. h. negativer bzw. weniger (je nach Vorzeichen), als die anderen Regelanwender.

Die Varianzanalyse zur Überprüfung der Effekte von Alter und prädiktiver Regelanwendung auf die Diskriminationen bei der eigenen Machbarkeit bei der unterschiedlichen Intensitätskombination ergab einen Trend bei der Wechselwirkung von Alter und prädiktiver konsistenter Regelanwendung, $F(7, 139) = 2.02$, $p = .057$. Die Ergebnisse sind in Abbildung 3.4.15 in Form eines Liniendiagramms dargestellt.

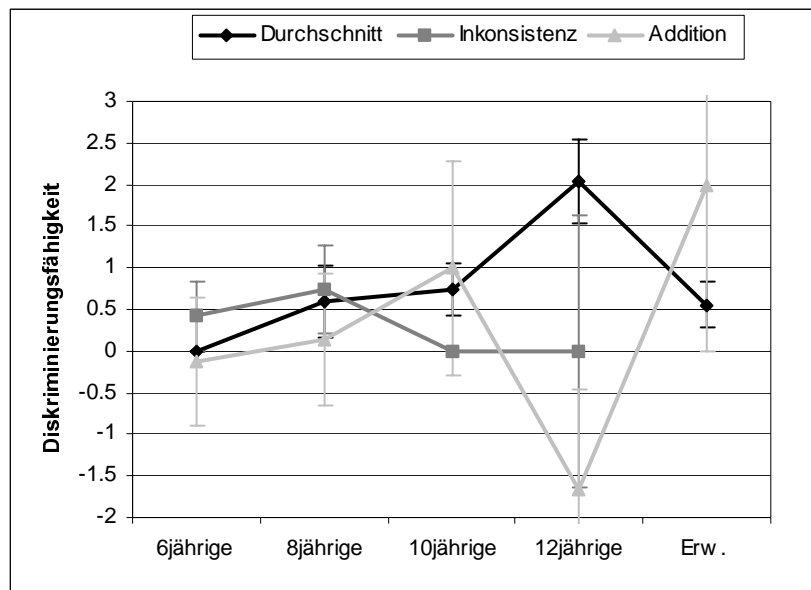


Abbildung 3.4.15. Mittelwerte und Standardfehler der Diskriminationen der eigenen Machbarkeitszuschreibungen bei der unterschiedlichen Intensitätskombination in Abhängigkeit vom Alter und der konsistenten prädiktiven Regelanwendung. Positive Diskriminationswerte besagen eine höhere Machbarkeitszuschreibung bei den physikalisch korrekten Mischungen und negative Werte höhere Machbarkeitszuschreibungen bei den physikalisch unkorrekten Mischungen.

Anhand der Abbildung wird deutlich, dass sich die unterschiedlichen Regelanwendungen in Phase 1 bei den Zwölfjährigen sehr stark auf die eigenen Machbarkeitseinschätzungen auswirkten: Während die Durchschnittsregelanwender bei dieser Altersgruppe ein Maximum im positiven Bereich erreichten, sind die falschen Anwender (sowohl Inkonsistente als auch Addierer) deutlich mit einem Maximum im negativen Bereich. Bei genauerer Betrachtung zeigte sich, dass zwei der drei zwölfjährigen Addierer deutlich negativ diskriminierten und die dritte Person gar nicht diskriminierte.

Exkurs: Skeptische Einstellung gegenüber der Fähigkeit des Zauberers, die Tricks reproduzieren zu können

Im Abschnitt zu den Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit von der Mischungsart (siehe S. 178ff. und insbesondere Abbildung 3.4.3) zeigten sich bei den Machbarkeitszuschreibungen zum Zauberer bei den Zehn- und Zwölfjährigen eine signifikante Unterscheidung zwischen den beiden Mischungsarten, welche bei den jüngeren Kindern nicht zu beobachten war. Bei den älteren Kindern war also eine bedeutsame Skepsis gegenüber den Fähigkeiten des Zauberers, den Farbmischungstrick reproduzieren zu können, zu beobachten. Da dieses Ergebnis theoretisch von grosser Bedeutung sein könnte, wird in diesem Kapitel den skeptisch urteilenden Kindern noch

genauer nachgegangen. Es werden Fragen beantwortet wie etwa, wie viele Kinder überhaupt skeptisch urteilten, zu welchem Ausmass diese Skepsis auftrat und ob sich diese skeptischen Kinder auch bei anderen Variablen von den nicht-skeptischen Kindern unterscheiden haben. Weiter interessierte auch, ob es sich bei den Skeptikern beim Farbmischungstrick um dieselben Versuchspersonen handelte, welche auch beim Wassertrick skeptisch gegenüber den Fähigkeiten des Zauberers urteilten, und wie sich diese allfällig „doppelten“ Skeptiker über den Altersverlauf hinweg verteilten.

Als Skeptiker wurden diejenigen Versuchspersonen definiert, welche dem Zauberer bei den Machbarkeitszuschreibungen bei den physikalisch unkorrekten Mischungen nicht das volle Vertrauen entgegengebracht hatten. Es wurden also alle Versuchspersonen, welche bei der Variablen „ufzauber“ oder „gfzauber“ einen Wert < 4 angegeben hatten, zu den Skeptikern gezählt. An dieser Stelle wurden die Werte für die beiden Intensitätskombinationen zusammen ausgewertet, wozu ein neues Datenfile mit insgesamt 280 Versuchspersonen (Variable: ugskeps) erstellt werden musste.

Wie viele Kinder urteilten skeptisch?

Nachdem im Abschnitt zu den Machbarkeitseinschätzungen verschiedener Akteure in Abhängigkeit von der Mischungsart (siehe S. 178ff.) nur Mittelwerte und darauf basierende Berechnungen angeführt wurden, wird hier genauer darauf eingegangen, wie viele der Versuchspersonen denn überhaupt skeptisch urteilten. In Abbildung 3.4.16 werden die prozentualen Anteile bezogen auf die Altersgruppe der skeptischen Versuchspersonen beim Farbmischungstrick dargestellt.

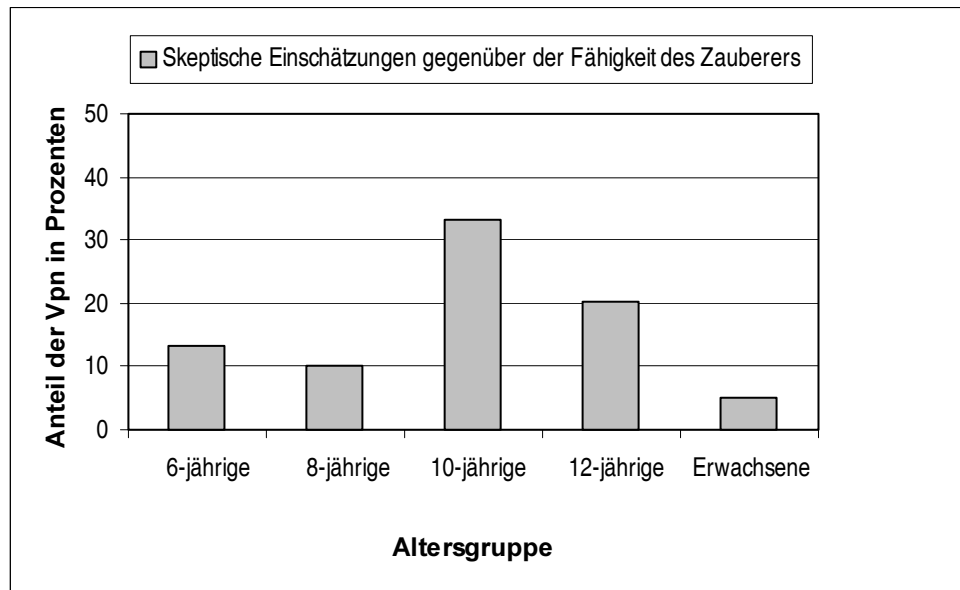


Abbildung 3.4.16. Prozentualer Anteil der Versuchspersonen in Prozent in Abhängigkeit von der Altersgruppe, welche gegenüber den Fähigkeiten des Zauberers, den Farbmischungstrick reproduzieren zu können, skeptisch eingestellt waren.

Die Überprüfung der unterschiedlichen Verteilung von skeptischen und nicht-skeptischen Versuchsteilnehmern über den Altersverlauf mit dem χ^2 -Test nach Pearson zeigte einen höchst signifikanten Unterschied, $\chi^2(4, n = 279) = 18.36; p = .001$. Zieht man die standardisierten Residuen zur Interpretation zur Hilfe, so zeigt sich, dass diese Signifikanz insbesondere durch die vielen Skeptiker bei den Zehnjährigen (33 Prozent der Altersgruppe bzw. 20 skeptische Urteile von insgesamt 60 Urteilen) sowie auch durch die wenigen Skeptiker unter den Erwachsenen (5 Prozent aller Erwachsene bzw. 2 von 40 Urteilen) und unter den Achtjährigen (10 Prozent der Achtjährigen bzw. 6 von 60 Urteilen) zu Stande kam.

Im Abschnitt zu den Machbarkeitsunterschieden beim Wassermischungstrick und den physikalisch korrekten Mischungen (S. 196ff.) zeigte sich ebenfalls eine Signifikanz bei den Zehn- und Zwölfjährigen bei der Prüfung mit den Mittelwerten der Zuschreibungen zum Zauberer. Da die Häufigkeiten bzw. Prozente der Skeptiker beim Wassertrick bei den Zehn- und Zwölfjährigen nicht deutlich höher waren als bei den anderen Altersgruppen, wurde an dieser Stelle auf die grafische Darstellung der prozentualen Anteile der Skeptiker verzichtet. Auch die Überprüfung der unterschiedlichen Verteilungen der skeptischen bzw. nicht-skeptischen Einschätzungen beim Wassertrick über den Altersverlauf mit dem χ^2 -Test nach Pearson ergab kein signifikantes Resultat. Die höchsten

Abweichungen zwischen beobachteter und erwarteter Anzahl wurden allerdings erneut bei den Zehnjährigen mit vielen Skeptikern (8 von 30 Urteilen bzw. 27 Prozent der Altersgruppe) und bei den Erwachsenen (1 von 20 Urteilen bzw. 5 Prozent aller Erwachsenen) mit wenigen Skeptikern beobachtet. Allerdings hatten beide standardisierten Residuen den Wert 1.3, was nach der Faustregel, nach welcher erst Werte > 2 eine Signifikanz anzeigen, aber zu tief für eine signifikante Abweichung ist (Bühl & Zöfel, 2002). Schliesslich wurden noch alle Skeptiker beim Wasser- und Farbmischungstrick zusammen genommen und ein neues Datenfile mit nun 420 Versuchspersonen erstellt (Variable: *ugwskeps*). Wie zu erwarten, ergab der χ^2 -Test nach Pearson mit den Verteilungen der skeptischen und nicht-skeptischen Versuchsteilnehmer beim Wasser- und Farbmischungstrick über den Altersverlauf hinweg eine höchst signifikante Unterscheidung, $\chi^2(4, n = 417) = 21.37; p < .001$. Die Abweichung der vielen skeptischen Urteile bei den Zehnjährigen (31 Prozent der Altersgruppe bzw. 28 von 90 Urteile) und die wenigen skeptischen Einschätzungen bei den Erwachsenen (5 Prozent bzw. 3 von 60 Einschätzungen) wiesen hohe standardisierte Residuen von 3.2 bzw. -2.3 auf, worin die Signifikanz begründet lag.

Wie skeptisch urteilten welche Kinder (Ausmass der Skepsis)?

Im Gegensatz zu den Berechnungen auf der Basis der Mittelwerte waren im obigen Abschnitt die signifikanten Unterschiede stets auf der Altersgruppe der Zehnjährigen begründet. Wie sich aber bei den varianzanalytischen Berechnungen zeigte, waren auch die Zwölfjährigen zu einem hohen Ausmass skeptisch gegenüber den Fähigkeiten des Zauberers, die Tricks reproduzieren zu können. Es liegt die Vermutung nahe, dass die offenbar selteneren zwölfjährigen Skeptiker in einem höheren Ausmass skeptisch waren als die jüngeren Kinder. Um diese Vermutung zu belegen, wurden die Skepsiswerte (d. h. die umkodierte Machbarkeitszuschreibungen zum Zauberer) von 0 bis 4 (keine Skepsis bis maximale Skepsis) addiert. Erneut wurden die Werte der gleichen als auch der unterschiedlichen Intensitätskombination zusammengefasst. Zudem wurden die Skepsiswerte auch beim Wassertrick addiert. Abbildung 3.4.17 zeigt die grafische Darstellung der addierten skeptischen Urteile bei den verschiedenen Mischungen in Abhängigkeit vom Alter.

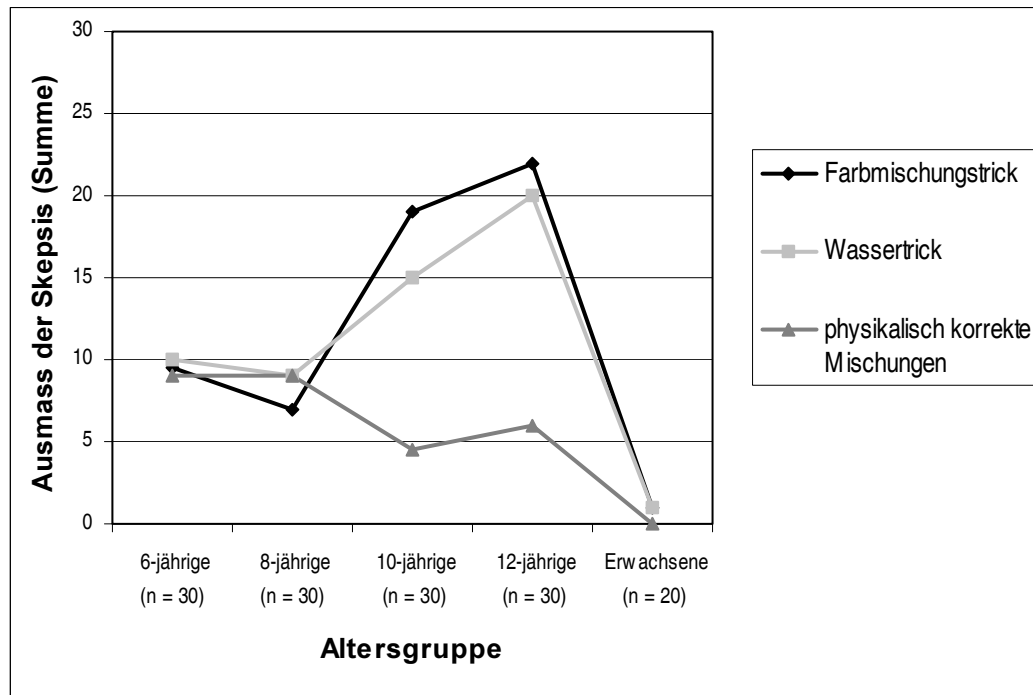


Abbildung 3.4.17. Ausmass der Skepsis gegenüber den Fähigkeiten des Zauberers, die verschiedenen Mischungsresultate reproduzieren zu können (addierte Skepsiswerte).

Die Abbildung zeigt bei den summierten Machbarkeitszuschreibungen zum Zauberer bei beiden Tricks ein Maximum bei den Zwölfjährigen, dicht gefolgt von den Zehnjährigen. Bedenkt man, dass diese skeptischen Werte aber nur auf wenige Zwölfjährige zurückgehen, dann müssen, diese wenigen Skeptiker in einem hohen Ausmass skeptisch gegenüber den Fähigkeiten des Zauberers sein. Dass sich diese Skepsis aber nur auf die Zuschreibungen der Fähigkeiten, die Tricks reproduzieren zu können, bezieht, zeigt der unterschiedliche Verlauf der Machbarkeitszuschreibungen bei den physikalisch korrekten Mischungen in der Abbildung oben. Die Unterschiede zwischen diesen Machbarkeitszuschreibungen der Tricks und den physikalisch korrekten Mischungen wurden schon oben auf ihre Signifikanz überprüft (siehe Tabelle 3.4.3 bzw. S. 196ff.). Es sei an dieser Stelle nochmals erwähnt, dass die Unterschiede bei den Zehn- und Zwölfjährigen bedeutsam waren.

Unterschieden sich Skeptiker und Nicht-Skeptiker hinsichtlich der Zuschreibung der eigenen Machbarkeit bei physikalisch korrekten und unkorrekten Mischungen?

Versucht man zu verstehen, weshalb diese Kinder den Fähigkeiten des Zauberers, die Tricks reproduzieren zu können, gegenüber so skeptisch waren, kann auch der Frage nachgegangen werden, ob diese Kinder sich selbst gegenüber bei den Tricks auch skeptischer eingestellt waren als die Nicht-Skeptiker.

Es wurden deshalb Varianzanalysen für die Zuschreibungen zur eigenen Machbarkeit mit dem Messwiederholungsfaktor Mischungsart (physikalisch korrekte und unkorrekte Mischungen) sowie den Zwischensubjektfaktoren Skepsis (auf der Basis der Definition von S. 203; = Einstellung gegenüber der Machbarkeit des Zauberers) und Alter gerechnet.

Die Varianzanalyse für die Zuschreibungen bei beiden Intensitätskombinationen zusammen zeigte nebst den erwarteten Haupteffekten der Mischungsart ($F(1, 269) = 41.61, p < .001$) und des Alters ($F(4, 269) = 4.76, p = .001$) auch einen Trend des Faktors Skepsis ($F(1, 269) = 2.98, p = .086$). Letzterer kam dadurch zu Stande, dass die Skeptiker bei beiden Mischungsarten zusammen tiefere Zuschreibungen machten. Von ganz besonderem Interesse war hier aber die Wechselwirkung zwischen den Faktoren Skepsis und Mischungsart, welche sehr signifikant war, $F(1, 269) = 7.89, p < .01$. Daneben war wie erwartet auch die Interaktion der Faktoren Mischungsart und Alter signifikant ($F(4, 269) = 4.84, p = .001$), was aber an dieser Stelle nicht von Bedeutung ist. Zudem wurden separate Varianzanalysen für die gleiche und unterschiedliche Intensitätskombination durchgeführt, deren Resultate im Wesentlichen der obigen Berechnung ähnlich waren²¹. In Abbildung 3.4.18 wird deutlich, wie sich die Faktoren Skepsis

²¹ Die Varianzanalyse den unterschiedlichen Intensitäten ergab Haupteffekte der Faktoren Mischungsart ($F(1, 130) = 26.18, p < .001$), Skepsis ($F(1, 130) = 4.48, p < .05$) und Alter ($F(4, 130) = 2.36, p = .057$). Auch die Wechselwirkungen der Faktoren Mischungsart und Skepsis sowie Mischungsart ($F(1, 130) = 10.22, p < .01$) und Alter ($F(4, 130) = 2.63, p < .05$) waren signifikant. Die varianzanalytische Berechnung bei den gleichen Intensitäten zeigte signifikante Haupteffekte der Faktoren Mischungsart ($F(1, 129) = 15.64, p < .001$) und Alter ($F(4, 129) = 2.56, p < .05$). Bei den Wechselwirkungen zeigten folgende bedeutsamen Resultate: Mischungsart und Alter ($F(4, 129) = 3.35, p < .05$) sowie Mischungsart, Alter und Skepsis ($F(4, 129) = 2.2, p = .072$).

und Mischungsart zusammen auf die Zuschreibungen zur eigenen Machbarkeit auswirkten.

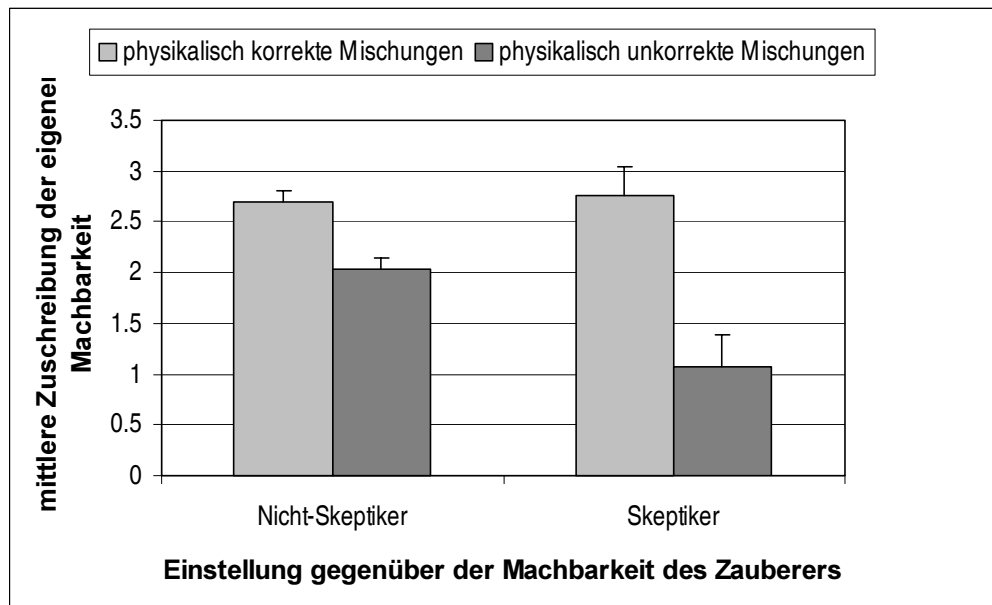


Abbildung 3.4.18 Zuschreibungen zur eigenen Machbarkeit, die Mischungen zu reproduzieren, bei beiden Intensitätskombinationen und allen Versuchspersonen in Abhängigkeit von der Mischungsart und der Einstellung gegenüber der Machbarkeit des Zauberers (Mittelwerte und Standardfehler).

Das Balkendiagramm zeigt, dass bei den Zuschreibungen zur eigenen Machbarkeit bei den physikalisch korrekten Mischungen kein Unterschied zwischen Skeptikern und Nicht-Skeptikern bestand. Dies bedeutet, dass sich diejenigen Versuchspersonen, welche dem Zauberer gegenüber beim Trick skeptisch waren, die korrekten Mischungen genauso zutrauen, wie die anderen Versuchspersonen. Ein Unterschied zwischen den Skeptikern und den Nicht-Skeptikern besteht aber bei der Zuschreibung zu sich selbst bei den physikalisch unkorrekten Mischungen. Die Skeptiker gegenüber dem Zauberer waren also auch bei sich selbst skeptischer als die übrigen nicht-skeptischen Versuchsteilnehmer.

Auch die Überprüfung der Unterschiede bei den Zuschreibungen zur eigenen Machbarkeit beim Wassertrick zwischen den Skeptikern und Nicht-Skeptikern mit der univariaten Varianzanalyse zeigte einen signifikanten Haupteffekt des Faktors Skepsis, $F(1, 137) = 6.2$, $p < .05$, sowie einen deutlichen Trend bei der Wechselwirkung zwischen den Faktoren Skepsis und Alter, $F(4, 137) = 2.41$, $p = .053$. Das von SPSS ausgegebene Diagramm zeigt, dass die zehn- und zwölfjährigen sowie die wenigen sechsjährigen und

erwachsenen Skeptiker sich selbst auch als weniger fähig einschätzten, den Wassertrick reproduzieren zu können. Bei den Achtjährigen waren aber die Skeptiker beim Zauberer sich selbst gegenüber weniger skeptisch, was die signifikante Wechselwirkung erklärt.

Verteilungen der „doppelten“ Skeptiker beim Farbmischungstrick und der Skeptiker beim Wassertrick

Zunächst wurde der Frage nachgegangen, ob zwischen den Skeptikern beim Farbmischungstrick und denjenigen beim Wassertrick ein Zusammenhang bestand. Dazu wurden die Verteilungen der beiden Variablen mit dem χ^2 -Test nach Pearson auf ihre Unabhängigkeit geprüft. Der höchst signifikante Prüfwert ($\chi^2(1, n = 275) = 40.61; p < .001$) lässt den Schluss zu, dass die beiden Variablen voneinander abhängig waren. Besonders auffällig war das sehr hohe standardisierte Residuum mit einem Wert von 5.3 bei den „doppelten“ Skeptikern. 50 Prozent (23 von 46 Urteilen) derjenigen, welche beim Farbmischungstrick die Fähigkeiten des Zauberers skeptisch beurteilten, waren auch beim Wassertrick skeptisch, d. h. sie waren doppelte Skeptiker. Abbildung 3.4.19 veranschaulicht, welchem Alter diese doppelten Skeptiker zuzuordnen waren.

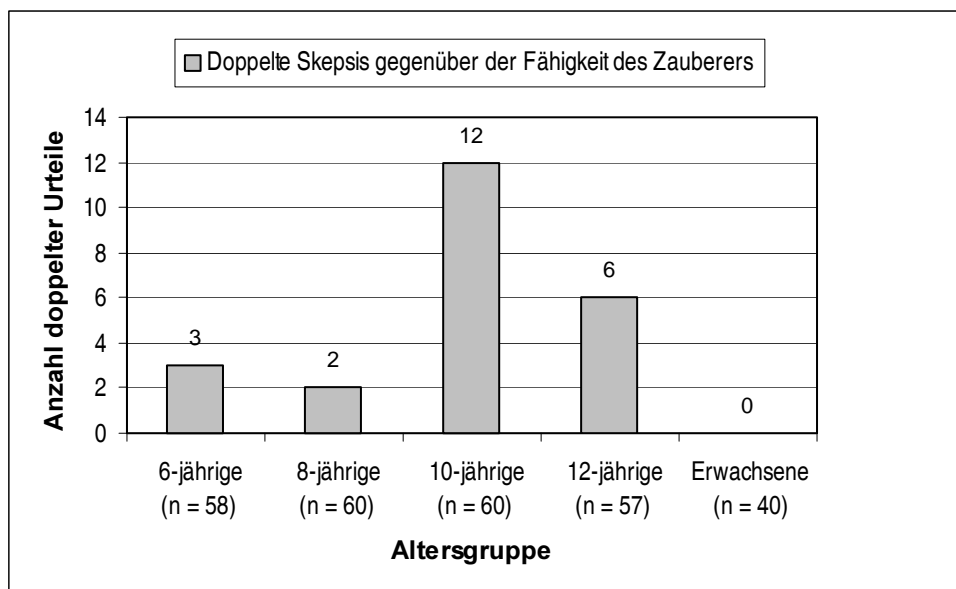


Abbildung 3.4.19. Häufigkeiten derjenigen Versuchspersonen, welche sich sowohl beim Wasser- als auch dem Farbmischungstrick skeptisch gegenüber den Fähigkeiten des Zauberers, die Tricks reproduzieren zu können, zeigten.

Passend zum obigen Balkendiagramm wurden die Verteilung der Variable „doppelte Skepsis“ mit dem χ^2 -Test nach Pearson auf ihren Zusammenhang mit dem Altersverlauf

überprüft. Es zeigte sich eine deutliche Signifikanz bei der Berechnung mit allen Versuchspersonen, $\chi^2(4, n = 275) = 17.35; p < .01$. Allerdings traten in 30 Prozent der Zellen der Kreuztabelle erwartete Häufigkeiten von < 5 auf, wonach die Voraussetzung des χ^2 -Tests knapp nicht erfüllt war. Es wurde deshalb eine erneute Berechnung ohne die Gruppe der Erwachsenen, welche keinen einzigen doppelten Skeptiker stellten, durchgeführt. Es zeigte sich immer noch ein sehr signifikantes Resultat ($\chi^2(3, n = 235) = 11.35; p = .01$), welches auf die vielen doppelten Skeptiker bei den Zehnjährigen (20 Prozent der Altersgruppe bzw. bei 12 von 60 Urteilen) zurückzuführen war.

Unterschieden sich die Skeptiker von den Nicht-Skeptikern hinsichtlich ihrer prädiktiven Regelanwendung?

Zur Beantwortung der Frage, ob sich die Skeptiker von den Nicht-Skeptikern auch in ihrer prädiktiven konsistenten Regelanwendung unterscheiden haben, wurden erneut χ^2 -Tests nach Pearson berechnet.

Bei der gleichen Intensitätskombination zeigte sich, dass die Skeptiker in der prädiktiven Regelanwendung schlechter waren als die Nicht-Skeptiker. Die Kreuztabelle der Verteilungen der verschiedenen Regelanwender (ohne die Anwender der Subtraktionsregel) auf die beiden Gruppen zeigt die grössten Abweichungen bei den Skeptikern: Viele urteilten inkonsistent (45 Prozent aller Skeptiker bei der gleichen Intensitätskombination), wenige wendeten die Durchschnittsregel (27 Prozent aller Skeptiker bei der gleichen Intensitätskombination) an. Der χ^2 -Test mit den drei verschiedenen Regelanwendergruppen war signifikant, $\chi^2(3, n = 135) = 8.62, p > .05$). Bei der unterschiedlichen Intensitätskombination hingegen dominierten bei beiden Gruppen deutlich die Durchschnittsregelanwender, der χ^2 -Test ergab keine Signifikanz.

Unterschieden sich die Skeptiker von den Nicht-Skeptikern hinsichtlich ihrer Diskriminierungsfähigkeit bei den Mischungsresultaten?

Da sich also die Skeptiker durch schlechtere prädiktive Regeln auszeichneten, liegt die Frage nahe, ob diese denn auch bei der Diskriminierung der physikalisch korrekten und unkorrekten Mischungsresultate anhand der Zauberskala schlechter abgeschnitten haben. Da diese abhängige Variable wie auch die Machbarkeitszuschreibungen, auf welchen die Kategorisierung in Skeptiker und Nicht-Skeptiker ja basiert, beide nach der

Demonstration der Mischungsvorgänge erhoben wurden, sind sie vergleichbarer als die Variablen, welche im obigen Abschnitt verglichen wurden.

Allerdings ergaben verschiedene Arten von Berechnungen (sowohl mit der proportional skalierten als auch mit der kategorisierten Diskriminationsvariablen; mit und ohne Erwachsene; für beide Intensitätskombinationen getrennt wie auch mit beiden zusammen) keinerlei Unterschiede in der Diskriminationsfähigkeit zwischen den Skeptikern und den Nicht-Skeptikern. Die beiden Gruppen unterschieden sich weder bei der gleichen noch bei der unterschiedlichen Intensitätskombination in ihrer Fähigkeit, die physikalisch korrekten von den unkorrekten Mischungsergebnissen zu unterscheiden.

Erklärungen zu den Entstehungsbedingungen der gezauberten Mischungen

Im Anschluss an die Präsentation aller Experimentaldurchgänge wurden die Kinder nach möglichen Ursachen der gezauberten Mischungen wie folgt befragt: „Hast du eine Ahnung oder eine Idee, wie ich das gemacht habe?“ Falls das Kind keine spontane Antwort gegeben hatte, wurde es direkt danach gefragt, ob es glaube, dass zum Zaubern richtige Zauberkraft nötig sei.

Die Auswertung der erhobenen Daten war allerdings nur im beschränkten Mass möglich, weil sich bei der Datenerhebung die folgenden zwei Probleme ergeben hatten:

1. Da die beiden Versuchsleiterinnen die Kinder unterschiedlich befragten, waren die Instruktionen offenbar zu ungenau formuliert gewesen. Während die eine Versuchsleiterin beim Ausbleiben einer Antwort auf die erste offene Frage, nicht weiter insistierte und die ausbleibende Antwort als „weiss nicht“ taxierte, fragte die andere weiter nach dem Glauben an Zauberkraft. Daraus ergab sich, dass die erste Versuchsleiterin sehr viel mehr Antworten in der Kategorie „weiss nicht“ erhalten hatte und eine Vergleichbarkeit mit der zweiten Versuchsleiterin nicht mehr gegeben war.
2. Ein weiteres Problem ergab sich bei der Datengewinnung, bei der die Erklärungen nur in schriftlicher Form von den Versuchsleiterinnen festgehalten wurden. Dadurch lag bereits eine Verzerrung der genannten Erklärungen vor, die den Standards der Wissenschaftlichkeit bezüglich Objektivität und Reliabilität nicht genügten. Allerdings muss hier angefügt werden, dass die Erfassung der kindlichen

Erklärungen zur Entstehung des Farbmischungstricks auch nicht im Vordergrund der Untersuchung stand und der Mehraufwand durch die mündliche Aufnahme der Erklärungen, anschliessender Transkription und Kodierung mit ausreichend hoher Interraterreliabilität nicht gerechtfertigt gewesen wäre. Zudem bestanden schon zu Beginn der Untersuchung Zweifel gegenüber einer direkten Erfassung des kindlichen Kausalitätsverständnisses in diesem Kontext. Das Hauptargument gegen eine direkte Erfassung bestand in der mangelnden Verbalisierungsfähigkeit der Kinder, was sich auch durch die wenigen spontanen Äusserungen der Kinder in der vorliegenden Untersuchung bestätigt hatte.

Da die erfassten Erklärungen nichts desto trotz das Bild der indirekten Erfassung durch die Machbarkeitszuschreibungen abrunden, wurden die Ergebnisse mit Einschränkungen ausgewertet: Es wurde auf statistische Berechnungen verzichtet und die Ergebnisse werden nur beschreibend dargestellt. Weiter wurden nur diejenigen Daten miteinander in Beziehung gesetzt, welche von der gleichen Versuchsleiterin erhoben wurden: Da die Acht- und Zwölfjährigen mehrheitlich von der Versuchsleiterin 1 befragt wurden, werden die Erklärungen dieser Altersgruppen gemeinsam grafisch dargestellt (Abbildung 3.4.20). Versuchsleiterin 2 erhob fast ausschliesslich die Daten der Sechs- und Zehnjährigen und der Erwachsenen, weshalb diese Gruppen miteinander verglichen werden können (Abbildung 3.4.21).

Die genannten Antworten wurden zunächst wie folgt kodiert:

a = Chemie / Mittel / Pulver / Farbe / Farbstoff

b = weiss nicht; keine Ahnung

c = richtig gezaubert / Zauberkraft

d = nicht richtig gezaubert / ohne Zauberkraft

e = Trick / Üben

f = eher nicht richtig gezaubert

g = Chemie, Zauberkraft nicht ausgeschlossen

h = Zauberkraft vielleicht

Im Anschluss an die Kodierung wurden die ursprünglichen Kategorien zu den fünf unten stehenden Kategorien zusammengefasst:

a => 1 = Chemie / Mittel / Pulver

d und e => 2 = ohne Zauberkraft / Trick

f, g und h => 3 = Zauberkraft nicht ausgeschlossen, aber unsicher

c => 4 = richtig gezaubert mit Zauberkraft

b => 5 = weiss nicht

Die Häufigkeiten dieser fünf Kategorien werden, für jede Versuchsleiterin getrennt, grafisch dargestellt.

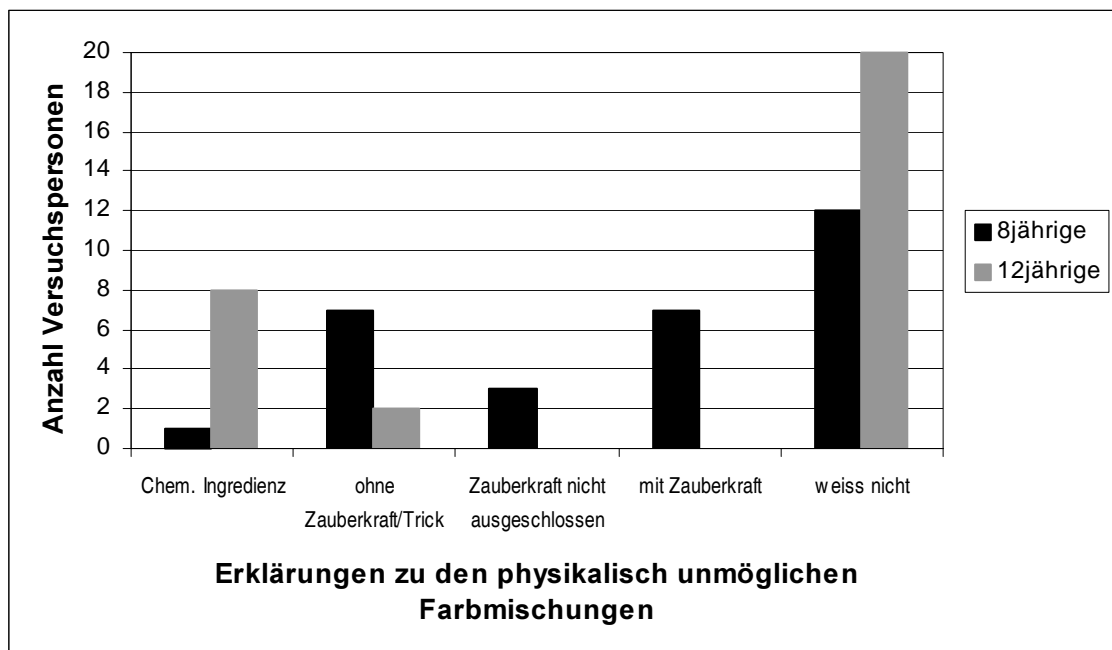


Abbildung 3.4.20 Häufigkeiten der verschiedenen Erklärungskategorien zur Entstehungsbedingung der physikalisch unmöglichen Mischungsart bei den Acht- und Zwölfjährigen, erhoben durch Versuchsleiterin 1.

Vergleicht man die Erklärungen zu den physikalisch unmöglichen Mischungsarten der Achtjährigen mit denjenigen der Zwölfjährigen, so fällt zunächst auf, dass nur Achtjährige den Einsatz von Zauberkraft vermuteten oder zumindest nicht ausschliessen konnten. Allerdings waren es auch mehr Achtjährige, die einen Trick ohne Einsatz von Zauberkraft in Erwägung zogen. Bereits 8 Zwölfjährige aber nur 1 achtjährige Versuchsperson meinten, dass eine chemische Ingredienz im Spiel war. Somit kann ausgesagt werden, dass auf der Basis dieser Versuchspersonen die Kinder mit zunehmendem Alter Zauberkraft ausschlossen, einen Trick vermuteten und sogar Einsicht in den tatsächli-

chen Entstehungsmechanismus der physikalisch unmöglichen Mischungsarten gewonnen hatten.

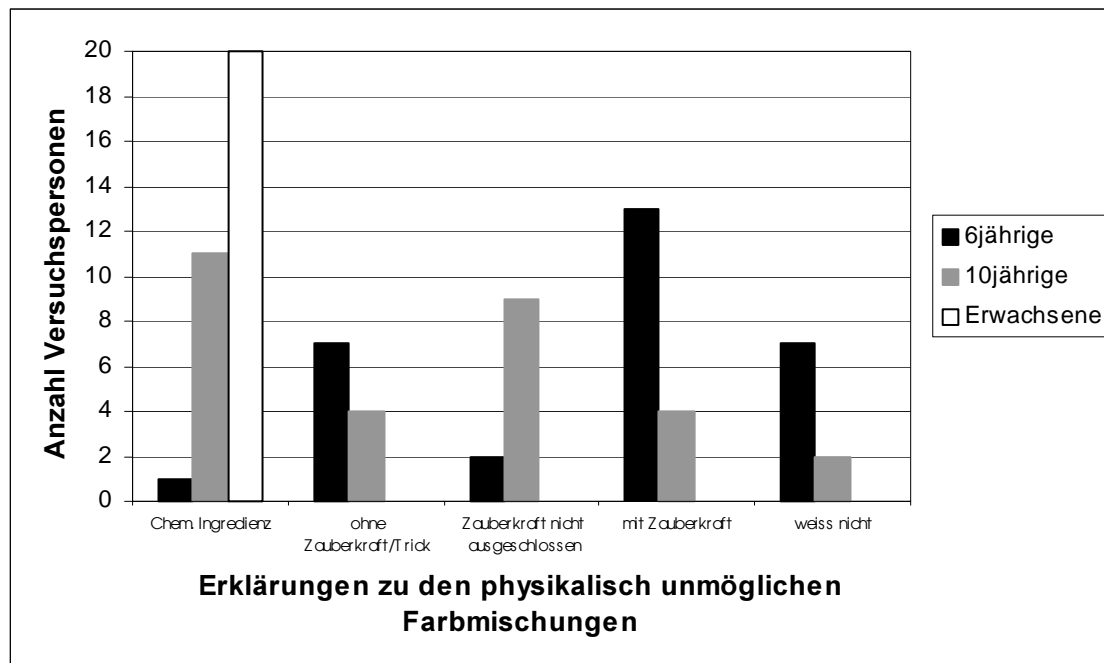


Abbildung 3.4.21. Häufigkeiten der verschiedenen Erklärungskategorien zur Entstehungsbedingung der physikalisch unmöglichen Mischungsart bei den Sechs- und Zehnjährigen sowie den Erwachsenen, erhoben durch Versuchsleiterin 2.

Bei den Erklärungen, welche durch Versuchsleiterin 2 erhoben wurden, zeigte sich wie schon bei der Versuchsdurchführung durch Versuchsleiterin 1 eine erhöhte Anzahl der Antworten der Kategorie „mit Zauberkraft“ bei den jüngeren Kindern sowie der Kategorie „chemische Ingredienz“ bei den älteren Kindern. Die Erwachsenen hatten alle ausnahmslos Einsicht in die Entstehungsbedingungen der physikalisch unmöglichen Mischungsart. Auch hier gaben die jüngeren Kinder häufiger als die älteren Kinder die Antwort, dass ein Trick dahinter stecke, ohne an Zauberkraft zu glauben. Immerhin 9 Zehnjährige waren zwar nicht sicher, ob nicht doch Zauberkraft im Einsatz war.

Zusammengefasst können die Entwicklungsverläufe der Erklärungen zur Entstehung der physikalisch unkorrekten Mischungsart folgendermassen interpretiert werden: Die Kinder gewannen zunehmend Einsicht in den Mechanismus des Farbmischungstricks. Ab dem Alter von zehn Jahren war der Anteil an Begründungen, die auf eine chemische Ingredienz hinwiesen, substantiell. Bei den Sechs- und Achtjährigen stand der Glaube an Zauberkraft im Zentrum. Sogar einige Zehnjährige konnten nicht ausschliessen, dass

beim Farbmischungstrick Zauberkraft im Spiel war. Jüngere Kinder, die nicht an Zauberkraft glaubten, vermuteten einen Trick hinter der Täuschung, hatten aber noch kaum eine Vermutung, wie dieser Trick zu Stande gekommen war.

3.4.5 Diskussion

Generelle Einflüsse der Mischungsarten und der verschiedenen Akteure auf die Machbarkeitszuschreibungen

Hinter der Erfassung der Zuschreibungen der Machbarkeit verschiedener Mischungsarten zu verschiedenen Akteuren stand die Hypothese, damit verschiedene Kausalbezüge von Kindern indirekt erheben zu können.

Die einfachste Hypothese bestand in einer erwarteten Unterscheidung der Kinder bei den Zuschreibungen zwischen den physikalisch möglichen und unmöglichen Mischungsarten. Diese Hypothese wurde klar bestätigt, da die physikalisch korrekten Mischungsarten bei allen Akteurzuschreibungen zusammen als deutlich machbarer eingeschätzt wurden als die physikalisch unkorrekten Mischungen. Diese Unterscheidung war – wie erwartet – vom Alter der Versuchspersonen abhängig: Die Sechsjährigen unterschieden die Mischungsarten gar nicht, bei den Acht- bis Zwölfjährigen wurde der Unterschied mit zunehmendem Alter immer grösser und die Erwachsenen unterschieden wieder weniger hoch. Dieser Effekt der Mischungsart entspricht der vorgängigen Hypothese, wird aber erst in Wechselwirkung mit anderen Faktoren inhaltlich besonders bedeutsam.

Hinsichtlich der Zuschreibungen zu unterschiedlichen Akteuren zeigten die Resultate, dass der Zauberer bei allen Versuchsdurchgängen zusammen als deutlich fähiger eingeschätzt wurde als die Lehrperson, die Eltern oder die Versuchsperson selbst. Zwischen Lehrperson und Versuchsperson selbst bestand aber auch ein deutlicher Unterschied: Die Lehrperson wurde als deutlich fähiger eingeschätzt als die Versuchsperson selbst. Kein grosser Unterschied zeigte sich aber zwischen den Machbarkeitszuschreibungen zu den Eltern und zur Versuchsperson selbst. Insgesamt scheint demnach die Figur des Zaubers mit anderen Attributen versehen zu sein als die anderen Personen. Diese Aussage erhält nach Betrachtung der Altersunterschiede innerhalb der Zuschreibungen der

einzelnen Akteure besondere Bedeutung: Während die Zuschreibungen bei den drei Akteuren Versuchsperson selbst, Eltern und Lehrperson sehr deutlich mit dem Alter höher wurden, veränderten sich diejenigen beim Zauberer nur gering. Während bei allen Altersgruppen die Kinder dem Zauberer deutlich mehr Machbarkeit zuschrieben als den anderen Akteuren, waren bei den Erwachsenen alle Akteure hoch. Angesichts des für erwachsene Studierende wenig komplexen Entstehungsmechanismus des Farbmischungstricks erstaunt dieses Resultat aber wenig.

Theoretisch besonders bedeutsam waren die Wechselwirkungen zwischen den Mischungsarten und den verschiedenen Akteuren bei den Machbarkeitszuschreibungen. Aufgrund der besonderen inhaltlichen Bedeutung der Zuschreibungen zur Versuchsperson selbst sowie zum Zauberer, werden diese getrennt diskutiert. Die Diskussion über die Zuschreibungen zum Zauberer erfolgt allerdings erst in einem späteren Abschnitt.

Exkurs: Machbarkeitszuschreibungen zur Versuchsperson selbst

Bei den Zuschreibungen zu sich selbst wurde erwartet, dass die physikalisch möglichen Mischungsarten mit zunehmendem Verständnis der kognitiven Durchschnittsbildung d. h. auch mit zunehmendem Alter als machbarer eingestuft werden. Gleichzeitig sollten sie auch höher eingestuft werden als die physikalisch unkorrekten Mischungsarten. Da aber auch hier vermutet wurde, dass die Kinder mit zunehmendem Alter mehr Einsicht in den chemischen Entstehungsmechanismus haben oder zumindest mit so genannten aussergewöhnlichen Transformationen vertrauter sind, war auch hier ein Anstieg der Machbarkeitszuschreibungen mit zunehmendem Alter zu erwarten. Diese zentralen Annahmen haben sich bei der vorliegenden Untersuchung bestätigt. Die Sechsjährigen, welche gemäss den Resultaten aus Phase 1 und 2 mehrheitlich noch nicht über die richtige Mischungsregel verfügten, unterschieden die beiden Mischungsresultate auch bei den Machbarkeitszuschreibungen nicht. Dies änderte sich ab dem Alter von acht Jahren: Die Kinder trauten sich selber auch die physikalisch korrekten Mischungen mit zunehmendem Alter deutlicher zu. Im Gegensatz dazu blieben die Machbarkeitszuschreibungen bei den physikalisch unmöglichen Mischungsresultaten bis zu den Zwölfjährigen tief und stiegen erst bei den Erwachsenen deutlich an. Es scheint als würde die Zuschreibung in die eigene Machbarkeit bei beiden Mischungsarten dem Kausalitätsverständnis zur Entstehung der beiden Mischungsarten entsprechen. Für eine Aussage über den generellen Entwicklungsverlauf der Einsicht in die Entstehungsbedingungen war es

nötig, nur die Daten derjenigen Versuchspersonen zu berücksichtigen, welche sowohl in Phase 1 als auch in Phase 2 im Sinne der korrekten Durchschnittsregel geantwortet hatten. Die Verläufe auf der Basis dieser Daten unterschieden sich insgesamt nur unwesentlich von den Verläufen basierend auf allen Versuchspersonen. Eindrücklich war aber trotzdem, wie stabil die Zuschreibungen bei der Lehrperson und der Versuchsperson selbst bezüglich der Machbarkeit der physikalisch unkorrekten Mischungsart in der ganzen Phase des Grundschulalters geblieben waren. Die postulierte Einsicht in den Entstehungsmechanismus des Farbmischungstricks war damit im Schulalter noch nicht gegeben.

Die genauere Betrachtung der Ausprägungen der Einschätzungen der eigenen Machbarkeitszuschreibung zeigte, dass der ausgebliebene Entwicklungsanstieg bei beiden Mischungsergebnissen bei den Zehnjährigen auf eine Übergangsphase zurückzuführen war: Bei beiden Mischungsarten nahmen die minimalen und maximalen Zuschreibungen bei den Zehnjährigen ab und die mittleren Machbarkeitszuschreibungen waren deutlich erhöht. Es sieht so aus, als wären die Zehnjährigen in ihrer Prognose vorsichtiger als die Zwölfjährigen, was zu tieferen Mittelwerten geführt hatte. Bei den eigenen Machbarkeitszuschreibungen zu den physikalisch unkorrekten Mischungen war über die verschiedenen Altersgruppen hinweg eine stetige Abnahme der minimalen Zuschreibungen zu beobachten. Im Gegensatz zum Verlauf bei den physikalisch korrekten Mischungsarten sind die minimalen Zuschreibungen aber deutlich verzögert. Das Gleiche lässt sich auch für die maximale Zuschreibung aussagen, da diese erst bei den Erwachsenen sehr häufig genannt wurde.

Die detaillierte Analyse der erhöhten mittleren Ausprägungen bei den Zehnjährigen ergab, dass diese nicht an einer alleinigen Erhöhung der neutralen Kategorie gelegen hatten. Die Vorgabe durch eine neutrale Kategorie „weiss nicht / unentschieden“ führte demnach nicht dazu, dass diese – aus welchen Motiven auch immer – mehr gewählt wurde als andere. Diese an sich generelle Problematik von neutralen Kategorien bei Rating-Skalen (Bortz & Döring, 2002, S. 216f.) scheint bei der verwendeten 5-stufigen Skala nicht zu inhaltlichen Schwierigkeiten beim Einschätzen durch die Kinder geführt zu haben.

Die Machbarkeitszuschreibungen zur Lehrperson verliefen über den Altersverlauf hinweg sehr ähnlich wie diejenigen zur Versuchsperson selbst, mit dem einzigen Unterschied, dass Erstere etwas höher waren als Letztere.

Einfluss der Intensitätskombinationen auf die Machbarkeitszuschreibungen

Die Machbarkeitseinschätzungen in Abhängigkeit der beiden Intensitätskombinationen zeigten, dass die Mischungen unterschiedlicher Intensitäten als leicht weniger machbar eingeschätzt wurden als die Mischungen gleicher Farbintensitäten. Hier zeigte sich aber noch eine dreifache Wechselwirkung zwischen Intensitätskombination, Akteur und Alter. Die Entwicklungsverläufe liessen am ehesten die folgende Zusammenfassung zu: Je älter die Kinder waren, umso „näher“ war die höhere Machbarkeit der Mischungen gleicher Intensitätskombination bei der Versuchsperson selbst. Den Zauberer kann man allerdings von den anderen Akteuren getrennt betrachten, denn ihm wurde sowieso mehr Fähigkeit zugeschrieben, wobei sich diese im Faktor Intensitätskombination nicht unterschieden hatte. Die Intensitätskombinationen wirkten sich aber bei den anderen Akteuren auf die Machbarkeitseinschätzungen aus. Die physikalisch korrekten Mischungen gleicher Farbintensitäten wurden bei den drei Akteuren als machbarer eingestuft als diejenigen Mischungen unterschiedlicher Farbintensitäten. Dies könnte an der höheren Vertrautheit mit der Mischung von Flüssigkeiten mit gleicher Farbintensität liegen, da die Einsicht in den physikalischen Entstehungsmechanismus ja mit zunehmendem Alter auch grösser wurde. Bei den unkorrekten Mischungsarten wurde aber bei den Akteuren Lehrperson und Eltern die Unterscheidung zwischen den beiden Intensitätskombinationen weniger deutlich oder gar nicht gemacht, während sie bei der Versuchsperson selbst noch gemacht wurde. Eine Interpretationsmöglichkeit wäre, dass die Kinder schon früher ahnen, dass möglicherweise hinter beiden Intensitätskombinationen der gleiche Entstehungsmechanismus steckt und deshalb andere Autoritätspersonen auch zu beiden gleich befähigt waren. Bei der Versuchsperson selbst hätte sich dies nicht gezeigt, weil sie aufgrund der eigenen Schwierigkeit, die unterschiedlichen korrekten Mischungen zu reproduzieren, auch bei den unkorrekten Mischungen mehr Schwierigkeiten bei der unterschiedlichen Intensitätskombination erwarten. Da diese Interpretation aber höchst spekulativ ist, können daraus kaum weitere Schlussfolgerungen gezogen werden.

Machbarkeitseinschätzungen beim Probeversuchsdurchgang „Wassermischungs-trick“ und Vergleich mit den beiden Mischungsarten des Experimentes

Die Resultate weisen darauf hin, dass die beiden Arten von Tricks hinsichtlich ihrer Machbarkeit nicht unterschieden wurden. Versteht man die Machbarkeitszuschreibungen als implizites Kausalitätsverständnis der Entstehungsbedingungen der verschiedenen Mischungsarten, so kann hieraus abgeleitet werden, dass die Kinder keinen Unterschied zwischen den beiden Trickarten bezüglich des Entstehungsmechanismus machten. Dieses Resultat ist im Hinblick auf die doch gemachte Unterscheidung bei der Zauberskala (vgl. weiter vorne) interessant. Sie schätzten den Wassertrick als mehr gezaubert ein als den Farbmischungstrick, dennoch wird erkannt, dass den beiden Trickarten der gleiche Entstehungsmechanismus zu Grunde liegt.

Kritisch muss allerdings noch angefügt werden, dass der Probeversuchsdurchgang des Wassertricks und die Experimentaldurchgänge der Farbtricks nicht in der Reihenfolge variiert wurden. Angesichts des nicht gefundenen Reihenfolgeeffektes innerhalb der Experimentaldurchgänge wäre aber auch kein solcher Effekt zu erwarten gewesen.

Diskriminationsunterschiede zwischen den verschiedenen prädiktiven Regelanwendern

Allfällige Unterschiede bei den Einschätzungen der verschiedenen Regelanwender aus Phase 1 wurden hauptsächlich bei den Diskriminierungen der beiden Mischungsarten bei der eigenen Machbarkeitszuschreibungen untersucht. Aufgrund der bisherigen Auswertungen wurde bei dieser abhängigen Variablen die Konfundierung mit der Einsicht in die unterschiedlichen Entstehungsmechanismen bei den Kindern ab acht Jahren als minimal eingeschätzt.

Zusammengefasst hing die Unterscheidung zwischen den physikalisch möglichen und unmöglichen Mischungsarten von der Regelanwendung in Phase 1 ab. Bei den Durchschnittsregelanwendern zeigte sich bei beiden Intensitätskombinationen der erwartete Anstieg mit dem Alter. Allerdings gelang den Acht- und Zehnjährigen die Diskriminierung bei der gleichen Intensitätskombination besser als bei der unterschiedlichen. Bei den inkonsistenten Regelanwendern aus Phase 1 wiederholte sich das Bild der Diskriminationen mit der Zauberskala: Bei der gleichen Intensitätskombination diskriminier-

ten sie zunehmend deutlicher, so dass sie schliesslich bei den Zehn- und Zwölfjährigen genauso deutlich diskriminierten wie die Durchschnittsregelanwender. Bei der unterschiedlichen Intensitätskombination war aber keine Verbesserung festzustellen. Bei den Addierern fielen die umgekehrten Einschätzungen der Mischungsergebnisse der unterschiedlichen Intensitätskombination bei den Zwölfjährigen auf: Sie unterscheiden genauso hoch wie die Durchschnittsregelanwender, nur in umgekehrter Richtung. Dieses Ergebnis kann im Sinne der Stabilitätshypothese als besonders stabile Additionsregel bei den älteren Kindern interpretiert werden und stützt auch hier die früheren Ergebnisse aus Phase 2. Bei der gleichen Intensitätskombination war aber keine Tendenz Richtung Stabilitätszunahme zu beobachten. Hier blieben die Einschätzungen relativ gleichförmig im indifferenten Bereich.

Machbarkeitszuschreibungen zum Zauberer

Während die Kinder ab dem Alter von 8 Jahren die physikalisch korrekten und nicht korrekten Mischungen bei den drei Akteuren Versuchsperson selbst, Eltern und Lehrperson unterschieden hatten, wurde diese Unterscheidung beim Zauberer nicht beobachtet. Sowohl bei den Sechs- als auch bei den Achtjährigen sowie den Erwachsenen war der Zauberer zu beiden Mischungen als gleich hoch befähigt eingestuft worden. Der Magier unterschied sich in seiner Fähigkeit zu zaubern eindeutig von anderen Autoritäten. Die Sechsjährigen konnten demnach zwar die physikalisch korrekten nicht von den unkorrekten Mischungen differenzieren, dennoch zeigten sie ein deutliches Verständnis vom Zaubern. Diese allumfassend hohen Einschätzungen des Zauberers änderten sich aber bei den älteren Kindern: Der Zauberer soll nicht mehr so selbstverständlich zur Herstellung der gezauberten Mischungen befähigt sein, wie es bei den jüngeren Kindern der Fall war. Bei den Zehn- und Zwölfjährigen trat eine ausgeprägte Skepsis gegenüber den Fähigkeiten des Zauberers auf. Auch im Vergleich mit den hohen Einschätzungen der Erwachsenen zeigten sich die älteren Kinder überkritisch gegenüber dem Zauberer.

Was lag dieser veränderten Einstellung gegenüber dem Zauberer zu Grunde? An einer allfälligen Einsicht in den Entstehungsmechanismus der gezauberten Mischungen konnte es nicht gelegen haben, da diese auch bei den Zwölfjährigen kaum gegeben war. Um zu klären, ob sich die skeptischen Kinder vielleicht auch bei anderen Einschätzungen in der vorliegenden Untersuchung von den nicht-skeptischen Kindern unterschieden hatten, wurden die Daten noch auf allfällige sonstige Unterschiede geprüft. Zudem wurde

versucht, mögliche Ursachen für die veränderte Einstellung gegenüber dem Zauberer bei den älteren Kindern zu eruieren.

Skeptische Einstellung gegenüber der Fähigkeit des Zauberers, die gezauberten Mischungen reproduzieren zu können

Am häufigsten skeptisch waren die Zehnjährigen, gefolgt von den Zwölfjährigen. Allerdings waren die Zwölfjährigen in stärkerem Ausmass skeptisch. Dies sowohl beim Farbmischungs- als auch beim Wassertrick. Die doppelten Skeptiker waren auch bei den Zehn- und Zwölfjährigen am häufigsten.

Die Abklärung, ob sich Skeptiker und Nicht-Skeptiker hinsichtlich der Zuschreibung der eigenen Machbarkeit bei richtigen und falschen Mischungen unterschieden hatten, ergab, dass die Skeptiker bei sich selbst bei den physikalisch unmöglichen Mischungsarten auch skeptischer waren. Interpretiert man die eigenen Machbarkeitszuschreibungen als indirekte Erfassung der Einsicht in den Entstehungsmechanismus, so könnte man vermuten, dass die Skeptiker die Entstehungsbedingungen der gezauberten Mischungen weniger verstanden hatten. Dies erklärt aber noch nicht, warum sie auch dem Zauberer die gezauberten Mischungen weniger zutrauten, als die nicht-gezauberten Mischungen.

Auch die Klärung der Frage, ob sich die Skeptiker von den Nicht-Skeptikern hinsichtlich ihrer prädiktiven Regelanwendung und der Diskriminierungsfähigkeit bei den Mischungsergebnissen unterschieden hatten, brachte keine annähernd ausreichende Erklärung für das skeptische Urteilsverhalten der Kinder. Es zeigte sich wohl bei den prädiktiven Mischungsurteilen der gleichen Intensitätskombination, dass die Skeptiker mehr addierten als die Nicht-Skeptiker. Da sich dieser Unterschied bei der Mischungsbeurteilung nach Beobachtung der Mischungsvorgänge aber nicht mehr zeigte, trägt dieses Ergebnis auch nicht viel zur Aufhellung der skeptischen Urteile auf.

Es fragt sich deshalb, ob die Zunahme der Skepsis mit dem einhergehenden Schwinden des magischen Denkens in Zusammenhang steht (siehe Ergebnisse zu den Erklärungen zu den Entstehungsbedingungen der gezauberten Mischungen, S. 211ff.). Magisches Denken wurde im theoretischen Teil (S. 43) bereits als magische Kausalitätskategorie definiert. Typisch für magisches Denken ist die Gewissheit, dass es bestimmte Personen

gibt, die über übernatürliche Kräfte verfügen und Naturgesetze tatsächlich durchbrechen können. Wäre es möglich, dass die Skeptiker dem Zauberer gegenüber besonders skeptisch waren, weil sie an seinen Zauberkräften zu zweifeln begannen? Wie sonst wäre es zu erklären, dass der Zauberer plötzlich nicht mehr zu den Tricks befähigt sein sollte?

Alleine das Ergebnis, dass bei den kindlichen Erklärungsversuchen zur Entstehung des Farbmischungstricks noch bei Zehnjährigen ein scheinbares magisches Denken vorhanden war, ist angesichts der vorliegenden Theorien (vgl. Kapitel 2.3 im Theorieteil, S. 39ff.) an sich schon erstaunlich.

Verschiedene Untersuchungen auf der verbalen Ebene zum magischen Denken von Kindern zeigten, dass Kinder ab vier Jahren kaum mehr spontane Bezüge zur Magie bei Erklärungsversuchen ungewöhnlicher Phänomene machten (Huang, 1930; Rosengren & Hickling, 1994; Rosengren et al., 1994). Wenn magische Erklärungen überhaupt gegeben wurden, dann kam dies nur vor, wenn keine anderen plausiblen kausalen Erklärungen zur Verfügung standen (Chandler & Lalonde, 1994; Phelps & Woolley, 1994). Diese Ergebnisse, die auf ein Auftauchen und Verschwinden des magischen Denkens im Vorschulalter schliessen liessen, standen Piagets Vorstellungen zu den magischen Jahren sehr nahe (Piaget, 1928; nach Johnson & Harris, 1994; Piaget 1930; nach Rosengren et al., 1994)²². Allerdings ist auch diese Sicht umstritten und wird schon seit längerem kontrovers diskutiert. Kritiker der magischen Jahre sind die Vertreter der *early competence-view* (z. B. Baillargeon, Spelke & Wasserman, 1985; Gelman & Baillargeon, 1983), welche sich für das Bild des Kindes als naiver Wissenschaftler aussprechen. Nach ihrer Sicht ist das kindliche Verständnis von physikalischen und biologischen Gesetzen zwar noch eingeschränkt, aber keinesfalls von magischer Kausalität durchsetzt.

²² An dieser Stelle scheint eine differenzierte Darstellung von Piagets Vorstellungen zum magischen Denken angebracht zu sein. Piaget konzentrierte sich bei seinen Untersuchungen mehr auf das magische Denken i. S. der besonderen Fähigkeiten der Kinder selbst, wie etwa die Überzeugung an die direkte Erfüllung eigener geäußelter oder gedachter Wünsche. Er grenzte diese Art von magischem Denken gegenüber so genannten sozialen magischen Überzeugungen ab (Piaget, 1929; nach Rosengren et al., 1994). Piaget nahm an, dass die sozialen magischen Überzeugungen systematischer und stabiler sind als die erst genannten. Weiter ging er davon aus, dass das magische Denken erst verstanden werden kann, wenn auch die sozialen magischen Überzeugungen geklärt sind. Diese Präzisierung zeigt demnach auf, dass Piagets Vorstellungen gar nicht im Widerspruch zum aktuellen Verständnis des magischen Denkens stehen.

Gegenargumente zu beiden Positionen kommen von verschiedenen Seiten: Beispielsweise ergab eine Umfrage bei Eltern vier- bis sechsjähriger Kinder (Rosengren et al., 1994, Studie 3), dass diese einen substanziellen Glauben an die Realität von Figuren wie Zauberer, Geister und Sankt Nikolaus bei ihren Kindern wahrnehmen würden.

Weitere Beweise lieferten die verschiedenen Untersuchungen von Subbotsky (z. B. 1985, 1994). Ein zentrales Ergebnis seiner Studien war, dass das magische Denken zwar auf der verbalen Ebene nicht geäußert wurde, dass aber auf der Handlungsebene der gleichen Kinder eine ausgeprägte Leichtgläubigkeit gegenüber magischen Fähigkeiten vorgefunden wurde. So wurde in einer aktuellen Untersuchung von Subbotsky (2001) leichtgläubiges Verhalten gegenüber der Zerstörungskraft eines Zauberspruches bei sechs- und neunjährigen Kindern beobachtet. Während die jüngeren Kindern auch in ihren verbalen Urteilen dem Zauberspruch hohe Wirkungskraft zubilligten, war dies bei den älteren Kindern nur auf der Handlungsebene zu registrieren. Auch in anderen Studien, die den Fokus auf das Verhalten der Kinder richteten, wurde eine deutliche Leichtgläubigkeit gegenüber der Magie vorgefunden (Harris, Brown, Marriot, Whittal & Harmer, 1991; Johnson & Harris, 1994; Woolley, 1997).

Fasst man die Ergebnisse verschiedener aktueller Studien zum magischen Denken bei Kindern zusammen, so scheint es dieses Denken wohl zu geben, dies allerdings in starker Abhängigkeit von der Untersuchungsmethode und -situation. Als Gründe für das nicht immer auftauchende magische Denken wurden genannt: Kindliche magische Erklärungen würden von Kindern nur selten geäußert, weil sie sozial unerwünscht seien (Woolley & Phelps, 1994). Zudem wird vermutet, dass die mangelnden Anzeichen eines magischen Denkens an der Methode der verbalen Erfassung kindlicher Erklärungen lagen (z. B. Subbotsky, 2001). Sowohl Rosengren und Hickling (1994) als auch Subbotsky (1985, 1994) betonten die Wichtigkeit eines magischen Kontextes zur Evokation magischer Kausalitätsbezüge. Die Präsentation der Versuchsleiterin als Zauberin in der vorliegenden Untersuchung könnte sich demnach entscheidend auf die häufigen Äußerungen der Kinder, dass Zauberkraft im Spiel war, ausgewirkt haben.

Neuere Untersuchungen zum kindlichen magischen Denken lassen demnach die erfassten magischen Erklärungen als plausibel erscheinen. Ob das schwindende magische

Denken allerdings mit der einhergehenden Skepsis gegenüber dem Zauberer im Zusammenhang steht, bleibt im Rahmen dieser Untersuchung spekulativ.

4 Allgemeine Diskussion

Der Ausgangspunkt der vorliegenden empirischen Untersuchungen bestand in der Untersuchung der kindlichen Vorstellung beim Zusammenfügen intensiver Grössen in veränderten Aufgabendarbietungen. Bisherige Untersuchungen zur kognitiven Durchschnittsbildung bei intensiven Grössen legten nahe, die Aufgabensituationen zu variieren und die kindlichen Lösungsstrategien in veränderten Aufgabenstellungen zu prüfen. Dies wurde in einem ersten Schritt durch eine geringe Änderung zu früheren Experimenten vollzogen, die – wie sich zeigte – bereits genügte, um alternative Lösungsstrategien zu provozieren. Die erste Phase diente weiter der Diagnostizierung der verwendeten Lösungsstrategien, um deren Stabilität und Grad der Reife in einer nächsten Phase zu überprüfen. Die neu modifizierte Methode des *magic experiment* in der zweiten Phase erlaubte aber nicht nur eine Aussage über die Beständigkeit und den Grad an Bewusstheit der Lösungsstrategien, sondern bot auch eine deutlich veränderte Aufgabensituation zur Einschätzung des zu Grunde liegenden physikalischen Prinzips. Im Rahmen der Schlussdiskussion werden die neuen Ergebnisse mit denjenigen früherer Studien zusammengefügt und es wird versucht, ein ganzes Bild der Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen zu zeichnen.

Ein weiterer Schwerpunkt der vorliegenden Arbeit lag in der Untersuchung des inhaltlichen Verständnisses gezauberter Ereignisse am Beispiel von physikalisch möglichen und unmöglichen Mischungen intensiver Grössen in einer dritten Phase. Es zeigte sich, dass die Methode des *magic experiment* dazu verwendet werden konnte, das kindliche Verständnis kausaler Entstehungsbedingungen physikalisch möglicher und unmöglicher Ereignisse zu untersuchen.

4.1 Theoretische Schlussfolgerungen zur Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen

Die Entwicklung der kognitiven Durchschnittsbildung bei intensiven Grössen wurde bis in die neunziger Jahre durch die Studien von Strauss und Mitarbeitern (z. B. Stavy et al., 1982; Strauss & Stavy, 1982) als u-förmiger Entwicklungsverlauf beschrieben. Allge-

mein formuliert, entsprach die dreiphasige Entwicklung einem Wechsel von einem intuitiven repräsentationalen System (Phase 1) über eine Übergangsphase (Phase 2) in ein späteres System numerischer Art (Phase 3). Die Übergangsphase der Kinder im frühen Schulalter war gekennzeichnet durch häufige additive Fehlvorstellungen.

Die Gruppe um Moore und Dixon (Ahl, Moore & Dixon, 1992; Dixon & Moore, 1996; Moore, Dixon & Haines, 1992) identifizierte in mehreren Studien vier Teilstrategien, die für ein perfektes Verständnis der kognitiven Durchschnittsbildung nötig sind. Ein annähernd vollständiges Verständnis aller Prinzipien wurde erst bei den 11. Klässlern und den Erwachsenen gefunden. Sie unterschieden zwischen intuitiven und analytischen Problemlösungsstrategien, die je nach Art der Aufgabendarbietung induziert würden. Ahl, Moore und Dixon (1992) postulieren zwei gleichzeitig ablaufende Entwicklungsprozesse: Zum einen entwickle sich mit zunehmendem Alter eine wachsende Fähigkeit, sich der Regeln bewusst zu werden und diese auf andere Inhalte zu übertragen. Zum anderen werde der Wechsel zwischen intuitivem und analytischem Denken zunehmend leichter vollzogen. Grundschulkindern weisen sie grössere Schwierigkeiten beim spontanen Wechsel von verschiedenartigen Aufgaben als älteren Kindern zu. Fehlerhafte Lösungsstrategien könnten bei jüngeren Schulkindern durch die verfehlte Anwendung analytischen Denkens – z. B. induziert durch eine quantitative Darbietung der relevanten Grösse – zu Stande kommen.

Die Untersuchungen von Jäger (1996, Jäger & Wilkening, 2001) zur kognitiven Entwicklung der Durchschnittsbildung bei intensiven Grössen brachten in diesem Zusammenhang die folgenden relevanten Neuerungen: Mit der Farbintensität wurde eine direkt wahrnehmbare intensive Grösse eingeführt, welche es ermöglichte, die intensive Grösse nicht mehr quantitativ zu präsentieren. Durch die Methode des funktionalen Messens im Rahmen der Informationsintegrationstheorie konnte nachgewiesen werden, dass die relevanten Grössen tatsächlich nach bestimmten Regeln miteinander verknüpft wurden. Obwohl in Studien zur sozialen Kognition gezeigt wurde, dass auch jüngere Kinder grundsätzlich zur Durchschnittsbildung befähigt sind, wurde die relativ häufig vorkommende Additionsregel als vereinfachende Lösungsstrategie und Vorläufer zur Durchschnittsregel bestimmt. Erklärt wird die Anwendung der Additionsregel mit der generellen Verfügbarkeit der Additionsregel in physikalischem Aufgabenkontext. Allerdings hängen die verfügbaren Regeln auch stark von der Aufgabendarbietung ab. Der

allgemein gültige u-förmige Entwicklungsverlauf wurde durch die Untersuchungen von Jäger infrage gestellt bzw. aus heutiger Sicht relativiert.

Die vielfältigen Ergebnisse legten nahe, die Aufgabendarbietungen zu variieren und die häufigsten Problemlösestrategien Durchschnittsbildung, Addition und Inkonsistenz in Abhängigkeit von zwei verschiedenen Aufgabenstellungen zu prüfen. Gleichzeitig war es möglich, durch die Erfassung von weiteren Merkmalen des Wissenserwerbs zusätzliche Hinweise zur Entwicklung verschiedener Problemlösestrategien bei Mischungsaufgaben intensiver Grössen zu erhalten.

Die vorliegende empirische Untersuchung verfolgte die folgenden zwei Ziele:

- Erfassung der Adaptivität des kindlichen Denkens und Identifizierung der lösungsrelevanten Aufgabenmerkmale.
- Entwicklungsverlauf der zwei häufigsten Problemlösestrategien Durchschnitts- und Additionsregeln hinsichtlich subjektiver Sicherheit, Stabilität und bewusstem Umgang nach Konfrontation mit widersprüchlichen Ereignissen.

4.1.1 Adaptive Problemlösung bei der Mischung intensiver Grössen

Der Ansatz des adaptiven Denkens von Kindern stellt die Lösungsstrategien in Abhängigkeit von der spezifischen Aufgabendarbietung. Frühere Untersuchungen zur kognitiven Durchschnittsbildung bei der Mischung von intensiven Grössen legten nahe, die z. T. recht unterschiedlichen Einschätzungen von Kindern als Ausdruck von adaptivem Denken zu verstehen. Die vorliegende Untersuchung baute auf dieser Annahme auf und lieferte zusätzliche Ergebnisse durch Variation der Aufgabendarbietung.

Als eines der zentralen Resultate der vorliegenden Untersuchung gelten denn auch die unterschiedlichen Einschätzungen von Mischungsergebnissen der intensiven Grösse Farbintensität in Abhängigkeit von der spezifischen Aufgabendarbietung. Es wurde schon vorgängig vermutet, dass die unterschiedlichen Präsentationen der Mischungsaufgaben verschiedene Lösungsstrategien induzierten. So ging es in der Folge darum, die lösungsrelevanten Hinweise der Aufgabendarbietungen zu identifizieren. Zudem wurde in dieser Studie erstmals gezeigt, dass Kinder die Mischungen gleicher und unterschiedlicher Intensitäten unterschiedlich behandeln. Dieses Resultat wurde durch eine weiterführen-

de Untersuchung mit Rekonstruktionsaufgaben von Schär (2004) gestützt: Bei Rekonstruktionsaufgaben mit gleichen und unterschiedlichen Intensitätskombinationen wurden bei der gleichen Intensitätskombination bei acht- bis zwölfjährigen Kindern substantiell mehr Additionsregeln identifiziert. Demnach enthalten bereits die verschiedenen Intensitätskombinationen Hinweise, die zu unterschiedlichen Lösungsversuchen führen.

Als generelles Muster bei der Entwicklung der kognitiven Durchschnittsbildung wird nach Betrachtung sämtlicher Studien zum Thema folgendes Bild entworfen: Adaptiv scheint das Denken v. a. bis zum Alter von zehn Jahren zu sein. Ab dem Alter von zwölf Jahren waren die Mischungsurteile relativ unbeeinflusst von der Aufgabenstellung. Dennoch wurden die Aufgaben noch nicht im gleichen Ausmass einheitlich gut gelöst wie bei den Erwachsenen.

Die theoretischen Schlussfolgerungen zur adaptiven Problemlösung bei Mischungsaufgaben intensiver Grössen erfolgten auf der Datenbasis der prädiktiven Urteile von früheren Studien und denjenigen aus Phase 1 des vorliegenden Experimentes, sowie in einem zweiten Schritt auch auf den Einschätzungen aus Phase 2. Das dahinter liegende Erklärungsmodell bestand in der Annahme einer Aktivierung eines kognitiven Schemas in Abhängigkeit von der Aufgabendarbietung sowie einer Problemlösung anhand von Hinweisreizen bei ausbleibender Aktivierung (vgl. Anderson & Wilkening, 1991). Diese Annahme wird auch von diSessa (1993) gestützt, der postuliert, dass sich intuitive „Theorien“ aus verschiedenen Schemata zusammensetzen, die dementsprechend auch wenig kohärent sind. Gewisse Schemata wie zum Beispiel das Balance-Schema würden sich sogar zur Verankerung der normativen Theorie eignen.

Zur Abhängigkeit von der Aufgabensituation in prädiktiven Urteilsaufgaben wurde folgender Schluss gezogen:

Wenn ein korrektes kognitives Schema bzw. Funktionswissen aufgrund einer geeigneten Aufgabendarbietung aktiviert werden konnte, bestanden bei den Kindern bis zehn Jahren viel weniger Schwierigkeiten, die Aufgaben korrekt zu lösen. Im Falle der Mischungen unterschiedlicher Intensitätskombinationen im vorliegenden Experiment schien eine solche geeignete Aufgabe zur Aktivierung eines Schemas vorgelegen zu haben. Es wurde vermutet, dass bei der Zugabe einer hellen Flüssigkeit zu einer dunkleren

Flüssigkeit das kognitive Schema des mentalen Ausbalancierens aktiviert werden konnte. Ähnlich geeignet schien die Aufgabe bei Jägers Experiment zur Mischung gleicher Intensitäten gewesen zu sein (Experiment 1, 1996). Hier wurde am ehesten ein aktiviertes Modell der Identitätserhaltung in Erwägung gezogen. Die Annahme einer Strategie der Identitätserhaltung wird durch die Untersuchung von Stavy und Strauss (1992) gestützt, bei der diese Denkstrategie auch bei jüngeren Kindern zu korrekten Lösungen geführt hatte. Dieses auch jüngeren Kindern verfügbare Modell könnte die mehrheitlich korrekten Einschätzungen der Sechsjährigen bei Jäger (1996) erklären. Das postulierte Modell des mentalen Ausbalancierens in der vorliegenden Untersuchung schien bei der Mehrheit der Sechsjährigen allerdings noch nicht verfügbar gewesen zu sein, da nur etwa 30 Prozent der jüngsten Kinder die Mischungsresultate korrekt voraussagen konnten. Die häufig inkonsistenten und additiven Urteile bei der gleichen Intensitätskombination wurden hingegen dahingehend interpretiert, dass die Aktivierung eines Schemas möglicherweise in dieser Aufgabenstellung blockiert war. Da dieser Aufgabendarbietung vermutlich keine dominanten Hinweisreize zu entnehmen waren, war der Anteil an inkonsistenten Antworten beträchtlich.

Die Vermutung, dass bei der Unmöglichkeit einer Aktivierung eines zur korrekten Lösung führenden Schemas Hinweisreize zur Problemlösung gesucht wurden, führte zu folgenden Schlussfolgerungen: Konnten deutliche Hinweisreize ausgemacht werden, wurden die intensiven Grössen vorwiegend addiert. Solche deutlichen Hinweisreize waren die quantitative Darbietung der intensiven Grösse, die Wahlaufgabe sowie das gleichzeitige Zusammenleeren, das eine Verdopplung induziert haben könnte (bei Jäger, 1996, unterschiedliche Intensitätskombination). Bei den jüngeren Kindern waren häufig hohe kognitive Anforderungen bei der Problemlösung limitierend, indem sie die Kinder überforderten und zu häufigen inkonsistenten Urteilen führten.

Die Ergebnisse aus den Beurteilungen der physikalisch korrekten und unkorrekten Mischungsresultate von vorgeführten Mischungsvorgängen im *magic experiment* verstärkten in erster Linie die These der hohen Adaptivität von Kindern im Alter zwischen acht und zehn Jahren. Allerdings wurde bei dieser Aufgabenstellung im Gegensatz zu den prädiktiven Aussagen keine aktive Generierung eines kognitiven Schemas angenommen. Stattdessen wurden die guten Einschätzungen bei den oben genannten Altersgruppen beim *magic experiment* eher auf ein passives Wiedererkennen und einer bestehen-

den hohen Vertrautheit der Aufgabe aus dem Alltag erklärt. Die grössere Schwierigkeit beim Erkennen der korrekten Mischungen der unterschiedlichen Intensitätskombination wurde darauf zurückgeführt, dass die resultierende mittlere Intensität in diesem Fall schwieriger nachzuvollziehen war als die gleich gebliebene Intensität bei der Mischung gleicher Intensitäten. Die Aufgabendarbietung der Mischung gleicher Intensitäten im *magic experiment* stellte für die acht- und zehnjährigen Kinder im Vergleich mit der Prädiktionsaufgabe erleichterte Bedingung dar, die sie zu Gunsten verbesserter Einschätzungen adaptiv nutzen konnten.

Die Ergebnisse der vorliegenden Untersuchung führten zu einer Relativierung und einer Differenzierung des Gültigkeitsanspruches der Annahme eines u-förmigen Entwicklungsverlaufs. Ein Wechsel von einer intuitiven Repräsentationsebene hin zu einer Repräsentation analytischer Art wird zwar weiterhin angenommen, allerdings schlägt sich dieser Wechsel längst nicht immer in vermehrten Fehlurteilen bei den Kindern im Alter von etwa zehn Jahren nieder. Liegt nämlich eine Aktivierung eines prägnanten korrekten kognitiven Schemas vor, so dominiert dieses auch bei den Zehnjährigen und führt entsprechend häufig zu korrekten Voraussagen der Mischungsergebnisse. Allerdings ist das intuitive Denken der Kinder bis zum Alter von etwa zehn Jahren äusserst abhängig von der Aufgabendarbietung. Zudem führen Aufgaben, die hohe kognitive Anforderungen an die Kinder stellen, bei jüngeren Kindern vermehrt zu inkonsistenten Urteilen.

4.1.2 Gewissheit und Stabilität der verschiedenen Lösungsstrategien im Entwicklungsverlauf

Nebst der Identifizierung von verschiedenen Lösungsstrategien bei der Mischung von intensiven Grössen in variierenden Aufgabensituationen waren folgende Aspekte beim Wissenserwerb von Interesse: Wie entwickelt sich die Stabilität der Lösungsstrategien unter verschiedenen Bedingungen und mit welcher Gewissheit werden sie vertreten? Diesbezüglich wurde von der Hypothese ausgegangen, dass sich die Stabilität und die Gewissheit einer Regel erst in einem Prozess über einen längeren Zeitraum hinweg entwickeln müssen, bis die Regel vollständig stabil ist und mit grosser Gewissheit vertreten werden kann.

Nebst dem Entwicklungsprozess beim Erwerb der korrekten Lösungsstrategie waren aber auch Unterschiede zwischen den korrekten Durchschnittsregelanwendern und den Anwendern der fehlerhaften Additionsregel von Interesse. Wurde die falsche Vorstellung genauso stabil vertreten wie die korrekte und wie entwickeln sich die beiden Vorstellungen?

Rolle der subjektiven Sicherheit beim Erwerb von Wissen

Bereits in der prädiktiven Bedingung war es möglich, mehr über die angewendeten Lösungsstrategien durch Erfassung der entsprechenden subjektiven Sicherheit zu erfahren. Die Gewissheit, mit der eine Regel vertreten wird, ist Entwicklungsveränderungen unterworfen und zeigt den Grad einer Reife einer kognitiven Vorstellung an. Nebst Altersunterschieden in der Sicherheit des Wissens waren aber auch Unterschiede in der Sicherheit zwischen Anwendern von korrekten und unkorrekten Lösungsstrategien denkbar.

Zur Bedeutung der subjektiven Sicherheit beim Erwerb von Wissen wurde schon im Rahmen der Theorie von Piaget geforscht (z. B. Miller, Brownell & Zukier, 1977; Miller, 1986; Robert, 1989). Die Unsicherheit von Wissen ist nach Piaget (1957, 1974/1980, 1975/1985; für eine Übersicht z. B. Beilin, 1983, 1989; Chapman, 1988) Ausdruck eines kognitiven Ungleichgewichts. Diesem Ungleichgewicht wird in der kognitiven Entwicklung eine wichtige Rolle beim Vorantreiben der Entwicklung bis

zum erzielten Gleichgewichtszustand zugeschrieben. Im Zustand des Ungleichgewichts werden noch alternative Vorstellungen in Erwägung gezogen, die dementsprechend nicht mit vollständiger Gewissheit vertreten werden können.

Auch Acredolo und O'Connor (1991) schreiben individuellen Unsicherheiten, die durch unausgedrücktes oder auch explizit geäußertes Bewusstsein von inneren Konflikten, Dissonanzen oder Widersprüchen hervorgerufen werden, eine hohe motivationale Bedeutung in der kognitiven Entwicklung zu. Sie vermuten, dass Unsicherheiten in einem lang andauernden Prozess ausgehalten werden müssen, bis genügend Informationen erworben werden konnten und der geistige Fortschritt an einem Punkt angelangt ist, an dem das Wissen mit vollständiger Sicherheit gutgeheissen werden kann. Die Entdeckung von konfliktreichen Modellen und der Versuch, diese zu vereinbaren, würden die Entwicklung vorantreiben bis zu einem notwendigen strukturellen Wechsel. Acredolo und O'Connor (1991) fokussieren sich deshalb in ihrer Forschung auf das Stadium, in dem alternative Erklärungen erkannt werden, und auf den Prozess, der zu einem Entschluss führt, die eine Erklärung anderen Vorstellungen vorzuziehen (vgl. z. B. Byrnes & Beilin, 1991). Die Äusserung von Unsicherheit eines Wissens deutet demnach auf einen inneren Konflikt und auf mit Zweifel behaftete Wissensbestände.

Die erfasste subjektive Sicherheit der Mischungsurteile in der prädiktiven Bedingung zeigte einen deutlichen Entwicklungsverlauf: Die subjektive Regelsicherheit stieg ab dem Alter von zehn Jahren deutlich an und war bei den Erwachsenen maximal hoch. Daraus lässt sich schliessen, dass die Mischungsregel bei den Kindern im Alter zwischen zehn und zwölf Jahren einen deutlichen Entwicklungsfortschritt durchläuft, indem sie mit grösserer Sicherheit vertreten werden kann. Wie aber schon von Acredolo und O'Connor (1991) vermutet, entwickelt sich der Prozess der Stabilisierung des korrekten Wissens über einen längeren Zeitraum hinweg und ist erst bei den Erwachsenen abgeschlossen.

Die Resultate bei den Acht- und Zehnjährigen zeigten, dass die achtjährigen Durchschnittsregelanwender bei beiden Intensitätskombinationen sicherer waren als die Zehnjährigen. Dieses Ergebnis liess vermuten, dass bei den Sicherheitseinschätzungen bei den jüngeren Kindern eine Konfundierung vorlag und diese deshalb nicht als Mass für die Reife einer Regel verstanden werden dürfen. Hohe Sicherheitseinschätzungen von

Kindern unter zehn Jahren wurden auch schon früher regelmässig beobachtet. Acredolo und O'Connor (1991) kamen zum Schluss, dass Unsicherheiten im Wissen überhaupt erst ab dem Alter von zehn Jahren geäussert werden können (*willingness to withhold judgment*). Jüngere Kinder würden ihre Lösungsstrategien mit erhöhter Gewissheit gutheissen, auch wenn in einer Aufgabe ungenügende Informationen zur Lösungsfindung vorhanden sind. Acredolo und Horobin (1987) bezeichnen dies mit *premature closure* und verdeutlichen damit, dass sie das Denken in diesem Alter für unreif halten. Die Akzeptanz, dass ein Mangel an gefestigtem Wissen vorhanden ist, könne erst im Alter ab zehn Jahren beobachtet werden. Sie vermuten den Grund hierfür in den hohen Anforderungen an die metakognitiven Fähigkeiten der Kinder. Dies bedeutet aber eine Einschränkung der Anwendung der Methode: Die Erfassung der subjektiven Sicherheit wäre erst bei Kindern ab zehn Jahren sinnvoll.

Die Erfassungen der subjektiven Regelsicherheiten zeigten nebst der zunehmenden Sicherheit mit dem Alter auch einen sehr deutlichen Unterschied – unabhängig von der *premature closure* – zwischen den verschiedenen Regelanwendern. Die inkonsistenten Regelanwender waren deutlich unsicherer als die konsistenten Regelanwender. Dieses Resultat stützt die Hypothese, dass Unsicherheit ein Mass für einen inneren Konflikt darstellt. Die höhere Sicherheit der korrekten Durchschnittsbilder gegenüber den Addierern zeigte, dass die Regel der Addierer nicht so ausgereift ist wie die korrekte Regel. Die Unterschiede in der Sicherheit der korrekten und unkorrekten Problemlösungsstrategie könnten durch verschieden starke Modelle oder Wissensbestände ausgelöst worden sein: Da vermutet wurde, dass nur bei fehlender Aktivierung eines korrekten Schemas addiert wird, setzt sich die doch noch verfügbare, aber falsche Lösungsstrategie vermutlich auch aus weniger kohärenten Wissensbeständen zusammen. So gesehen, wäre die Sicherheit auch ein Mass für die Kohärenz der aktivierten Wissensbestände und deren Passung mit der Aufgabendarbietung. Würde diese Vermutung zutreffen, so müssten die Additionsregeln bei den quantitativ dargebotenen Aufgaben wie bei Strauss und Mitarbeitern (z. B. Stavy et al., 1982; Strauss & Stavy, 1982) mit grösserer Sicherheit vertreten werden als sie es bei dieser Aufgabe wurden. Dabei müsste aber beachtet werden, dass bei jüngeren Kindern eine *premature closure* die Einschätzungen beeinflussen könnte. Allerdings waren in der vorliegenden Untersuchung die Unterschiede in der Sicherheit von verschiedenen Wissensbeständen auch bei den jüngeren Kindern noch zu beobachten.

Die Entwicklungsverläufe der Durchschnittsregelanwender zeigten bei beiden Intensitätskombinationen einen Sicherheitseinbruch bei den zehnjährigen Kindern. Folgt man den Hypothesen von *willingness to withhold judgment* und *premature closure*, so können daraus folgende Annahmen geäußert werden: Kinder im Alter von zehn Jahren befinden sich in einer Phase von konfliktreichen Modellen, denen eine hohe motivationale Bedeutung in der geistigen Entwicklung zugeschrieben wird. Alternative Modelle sind in diesem Alter noch denkbar. Ein Entwicklungsfortschritt zwischen dem Alter von zehn und zwölf Jahren war anhand der Einschätzungen der Zwölfjährigen sehr deutlich zu beobachten. Jüngere Kinder hingegen urteilen auch auf der Basis von ungenügenden Informationen und sind sich dessen kaum bewusst. Eine starke Orientierung bei der Wahl von Lösungsstrategien an verfügbaren Hinweisreizen scheint plausibel und würde für ein intuitives Problemlösen sprechen.

Stabilität der verschiedenen prädiktiven Mischungsregeln nach Konfrontation mit widersprüchlichen Ereignissen

Miller (1976) hatte trotz der vielen methodischen Mängel von früheren Untersuchungen mit dem *extinction paradigm* an dieser Methode festgehalten und die dadurch fast einzigartige mögliche Erforschung der „Tiefe“ von Wissen (*in-the-system-knowledge*) akzentuiert. In der vorliegenden Untersuchung wurde versucht, die offensichtlichen methodischen Mängel zu beheben und die kindlichen Reaktionen auf physikalisch mögliche und unmögliche Mischungsergebnisse durch Realitätseinschätzungen anhand der Zauberskala zu erfassen. Durch diese indirekte Erfassung der Verwerfung bzw. des Festhaltens der ursprünglichen Mischungsregel war es möglich, auch jüngeren Kindern eine adäquate Urteilsskala in die Hand zu geben.

Die Resultate der Realitätseinschätzungen aus Phase 2 wurden bereits unter dem Aspekt des adaptiven Denkens berücksichtigt und diskutiert. Es bleibt deshalb in diesem Abschnitt, dieselben Resultate unter dem Aspekt der Stabilität der beiden häufigsten Lösungsvorstellungen der Durchschnitts- und Additionseignen zu diskutieren.

Ein zentrales Ergebnis mit der Methode des *magic experiment* bestand darin, dass sich die Durchschnittsregel als viel stabiler als die Additionsregel erwiesen hatte. Dasselbe zeigte sich auch bei den erhobenen Machbarkeitszuschreibungen in Phase 3. Die Vorstellung einer Intensivierung der Farbintensitäten nach Zusammenleeren von zwei Flüss-

sigkeiten aus der prädiktiven Bedingung wurde demnach nach Betrachtung verschiedener Mischungsvorgänge häufiger verworfen als die Vorstellung einer Durchschnittsbildung der beiden ursprünglichen Intensitäten.

Dieses Resultat ist insbesondere im Hinblick auf die Forschungsdiskussion zu den sog. Misskonzepten in der intuitiven Physik von Interesse. Man weiss heute aus vielen Untersuchungen, dass viele fehlerhafte physikalische Vorstellungen erstaunlich resistent gegenüber Instruktionen und anderen täglichen Erfahrungen sind, manchmal sogar bis ins Erwachsenenalter (siehe z. B. Carey, 1986; Clement, 1982; Halloun & Hestenes, 1985a, 1985b; McCloskey, 1983; Novak, 1987). Intuitive kognitive Fehlvorstellungen halten sich demnach trotz regelmässiger korrekter Rückmeldung durch die Natur hartnäckig bis ins Erwachsenenalter. Physikalische Misskonzepte stellen im Forschungskontext der kognitiven Entwicklungspsychologie ein geeignetes Forschungsfeld dar, um mehr über die generelle Entwicklung des Denkens und des Wissenserwerbs und allfälligen kognitiven Beschränkungen zu erfahren (vgl. Wilkening & Lamsfuss, 1993). Ein recht genereller Vorschlag, wie die Resistenz kindlicher physikalischer Fehlvorstellungen erklärt werden könnte, stammt von Sodian (1998): Die Fehlvorstellungen sind in alternative intuitive Theorien über physikalische Phänomene eingebettet und diese könnten erst aufgegeben werden, wenn das Begriffssystem und der Erklärungsapparat der intuitiven Theorien grundlegend verändert und neu aufgebaut wurden. Die vorliegenden Resultate legen allerdings die Vermutung nahe, dass es sich bei der additiven Fehlvorstellung im Gegensatz zur korrekten Vorstellung der Durchschnittsbildung eher um weniger verankertes Wissen handelt.

Die vorhandenen Befunde zur additiven Fehlvorstellung bei der Mischung intensiver Grössen wiesen auf eine deutlich geringe Stabilität hin. Von einem Misskonzept kann in diesem Fall kaum gesprochen werden. Der bewusste Umgang mit der Regel (vgl. weiter unten) wies allerdings darauf hin, dass die falsche Vorstellung von den Kindern nicht explizit verworfen werden konnte. Möglicherweise ist durch die Erfassung mit dem *magic experiment* tatsächlich eine fast einzigartige Methode gegeben, welche die Konflikthaftigkeit der physikalischen Fehlvorstellungen von jüngeren Kindern aufzudecken vermag. Die beobachtete Tendenz zur Stabilisierung der falschen Vorstellung bei den zwölfjährigen Kindern lässt allerdings vermuten, dass sie mit zunehmendem Alter tiefer verankert sind als bei jüngeren Kindern. Bei den erwachsenen Studierenden hingegen

wurde die selten angewendete Additionsregel im Verlaufe des Experimentes widerrufen. Ob die Daten der gut ausgebildeten Studierenden einen generellen Entwicklungstrend widerspiegeln, bleibt aber fraglich.

Die theoretische Annahme einer Stabilisierung der korrekten Durchschnittsregel mit zunehmendem Alter war klar. Konkret wendeten die Kinder ab dem Alter von acht Jahren die Durchschnittsregel recht stabil an. Die Durchschnittsregel nahm nach dem zwölften Lebensjahr aber noch an Stabilität zu, wie die erwachsenen Studierenden zeigten, welche die korrekte Regel bei beiden abhängigen Variablen fast maximal bestätigten.

Bewusster Umgang mit der Mischungsregel

Nebst der Realitätseinschätzung physikalisch korrekter und unkorrekter Mischungsergebnisse im *magic experiment* wurde auch auf den bewussten Umgang mit der ursprünglich geäußerten Mischungsregel abgezielt. Es zeigte sich aber schon bei der Versuchsdurchführung, dass die jüngsten Kinder, die Sechsjährigen, mit der Fragestellung überfordert waren. Der bewusste Umgang mit einer Regel stellte hohe Anforderungen an die Kinder bezüglich ihrer metakognitiven Fähigkeiten und war deshalb für jüngere Kinder keine geeignete Methode. Daraus liess sich aber auch schliessen, dass ein bewusster Umgang erst im Schulalter möglich ist.

Im Entwicklungsverlauf zeigten die *Durchschnittsregelanwender* nach Beobachtung der physikalisch korrekten Mischungsergebnisse einen stetigen Anstieg der Regelsicherheit auf einem hohen Niveau. Auch nach Beobachtung der physikalisch unkorrekten Mischungsergebnisse war tendenziell eine Zunahme der Sicherheit zu beobachten, welche aber auf einem deutlich tieferen Sicherheitsniveau ansetzte. Auch bei der konsistenten bewussten Bestätigung der Regel zeigte sich tendenziell Zunahme mit dem Alter, allerdings war bei den Achtjährigen noch fast kein konsistent korrekter Umgang zu beobachten. Zusammengefasst war auf der bewussten Ebene ein Entwicklungssprung mit deutlicher Zunahme an Stabilität zwischen acht und zehn Jahren zu beobachten. Dies entspricht der Hypothese von Acredolo und Horobin (1987), dass metakognitive Fähigkeiten erst im Alter von zehn Jahren im Umgang mit Wissensbeständen beobachtet werden können. Bei der Interpretation des Entwicklungsverlaufs der Regelbestätigung der *Addierer* wurde die hohe Bestätigung der Achtjährigen im Sinne des *premature closure* von Acredolo und Horobin (1987) interpretiert. Bei den Kindern von zehn bis zwölf Jahren war eine Tendenz zur Beibehaltung bzw. Bestätigung ihrer unkorrekten Regel festzustellen. Eine mögliche Erklärung dieser Tendenz wäre, dass die Zwölfjährigen den Widerspruch eher bemerken als die Zehnjährigen und sich eher gezwungen fühlen, bei ihrer früheren Aussage zu bleiben. Das Denken jüngerer Kinder schien offener gewesen zu sein und bei den älteren Kindern bestand eine grössere Tendenz, in verschiedenen Aufgaben die gleiche Regel anzuwenden.

Die Ergebnisse des *magic experiment* zum generellen Vergleich zwischen Durchschnitts- und Additionsregel bestärkten die Vermutung, dass die Addierer einem grösseren

ren inneren Konflikt ausgesetzt waren als die Durchschnittsregelanwender: Obwohl das nicht erwartete Mischungsresultat nicht nur bei den Additionsregelanwendern sondern auch bei den Durchschnittsbildern zu mehr Unsicherheit bzw. zu grösserem Konflikt führte als das erwartete Resultat, waren die Addierer auf deutlich tieferem Sicherheitsniveau.

Postulierter Entwicklungsverlauf nach Betrachtung der Resultate

Aus den Ergebnissen der Phasen 1 und 2 der vorliegenden Untersuchung wird folgender allgemeiner Entwicklungsverlauf der kognitiven Durchschnittsbildung skizziert:

- Sechsjährige: Bei den Vorschulkindern konnte keine dominante Mischungsregel ausgemacht werden. Die vorliegenden Aufgabenstellungen enthielten keine salienten Hinweisreize, die sie für ein verfügbares kognitives Schema nutzen konnten. Offen bleibt, ob das bei den älteren Kindern postulierte kognitive Schema des mentalen Ausbalancierens den Sechsjährigen überhaupt schon zur Verfügung steht. Für einen bewussten Umgang mit den Mischungsregeln scheinen die dafür notwendigen metakognitiven Fähigkeiten den Sechsjährigen noch zu fehlen.
- Achtjährige: Die Einschätzungen der Achtjährigen in verschiedenen Bedingungen wiesen auf eine deutliche Adaptivität an die Aufgabensituation hin. Die Hinweisreize der verschiedenen Aufgaben führten nicht in allen Bedingungen zur gleichen Aktivierung von kognitiven Schemata. Offen bleibt, ob in den verschiedenen Bedingungen jeweils andere Modelle bei der Lösungsfindung genutzt wurden oder ob die kognitive Kapazität nicht dazu ausreichte, das korrekte kognitive Schema bei allen Aufgaben konsistent anzuwenden. Beim bewussten Umgang mit der Regel zeigten sich auch bei den Achtjährigen Schwierigkeiten, die auf wenig metakognitive Einsicht in die eigene physikalische Vorstellung hinwiesen.
- Zehnjährige: Bei den Zehnjährigen war das adaptive Denken besonders ausgeprägt. Gleichzeitig drückten sie über die subjektive Sicherheit in der prädiktiven Bedingung aus, dass ihr Wissen noch konfliktreich ist. Dass ein bewusster Umgang mit der Regel bei den Zehnjährigen möglich wird, zeigte auch der bewusste Umgang mit der Regel nach der Präsentation verschiedener Mischungsresultate im *magic experiment*.
- Zwölfjährige: Die zwölfjährigen Kinder wendeten in höherem Ausmass als die jüngeren Kinder in verschiedenen Bedingungen die gleiche korrekte Regel an. Da

die Häufigkeiten korrekter Urteile in allen Bedingungen ähnlich hoch waren, war auch die Adaptivität des Denkens nicht mehr so ausgeprägt wie bei den jüngeren Kindern. Es fragt sich, ob in den verschiedenen Bedingungen dasselbe kognitive Modell zur Verfügung stand und sie dieses aufgrund höherer kognitiver Fähigkeiten auch unter erschwerten Bedingungen anwenden konnten. Denkbar ist theoretisch auch ein leichter Wechsel zwischen verschiedenen Modellen. Die hohen subjektiven Sicherheiten der prädiktiven Mischungsregeln wiesen daraufhin, dass ihr Wissen weniger konfliktreich geworden ist.

- Erwachsene: Die erwachsenen Studierenden urteilten fast ausnahmslos und in allen Bedingungen im normativen Sinne. Da auch die Stabilität und Gewissheit der Regeln sehr hoch waren, kann bei ihnen am ehesten von einem ausgereiften Konzept der kognitiven Durchschnittsbildung gesprochen werden.

4.1.3 Ausblick auf weiterführende Untersuchungen

Durch die vorliegende Untersuchung zeigte sich, dass ab dem Alter von zwölf Jahren ein bewussterer Umgang mit dem kognitiven Modell der Durchschnittsbildung möglich wird. Bei jüngeren Kindern wurde vermutet, dass aufgabenspezifisch kognitive Modelle zur Problemlösung aktiviert wurden. Es stellt sich nun die Frage, welche mentalen Modelle denn hinter den erfassten Verknüpfungsregeln der Durchschnittsbildung und Addition stehen. Beispielsweise war nicht klar, ob es immer dasselbe kognitive Modell war, das die Versuchspersonen zur korrekten Lösung führte.

Die postulierten unterschiedlichen Schemata sollten deshalb in Nachfolgeexperimenten geprüft werden. Eine Möglichkeit, die sich z. B. bei Lamsfuss (1994) auch bei jüngeren Kindern als sehr nützlich erwiesen hatte, besteht in der Aufnahme von Explorationsprotokollen. Sie bekam dadurch deutliche Hinweise darauf, dass der dort präsentierte Versuchsaufbau zur Addition von Kräfte-Vektoren mit einer Balkenwaage verglichen wurde. Nachfolgeexperimente von Lamsfuss (1994) konnten die Bedeutung der Balkenwaage-Analogie für das Urteilsverhalten bestätigen. Es wäre auch denkbar mit vorgängigen Übungsdurchgängen, die die Verwendung unterschiedlicher Lösungsstrategien anregen, den Einfluss von korrekten und falschen Analogien auf das Problemlöseverhalten zu überprüfen. Die grosse Bedeutung von Analogien bei der Bearbeitung physikalischer Aufgaben wurde sowohl für Kinder als auch Erwachsene mehrfach nachgewiesen

(z. B. Clement, 1990; Spiro, Feltovich, Coulsen & Anderson, 1989). Denkbar wären auch gezielte Prüfungen der differenziellen Einflüsse Paarvergleichsmethode vs. abgestufte Urteilsskalen sowie das gleichzeitige Zusammenleeren in drittes Glas vs. die Zugabe einer zweiten Flüssigkeit zu einer ersten Flüssigkeit jeweils in Abhängigkeit von der Intensitätskombination auf das Urteilsverhalten der Kinder.

Die aussagekräftigen Resultate der vorliegenden Untersuchung könnten durch weitere Experimente mit anderen intensiven Grössen bestätigt werden. Denkbar wären Aufgaben mit der Methode des *magic experiment* mit den intensiven Grössen Temperatur oder Süsse. Eigene Abklärungen hinsichtlich der chemischen Machbarkeit zeigten, dass chemische Manipulationen mit der Süsse sich eher für den Umgang mit Kindern eignen würden. Bei diesen intensiven Grössen müsste aber vorgängig noch die Wahrnehmungsfähigkeit von Intensitätsunterschieden bei qualitativer Darbietung bei Kindern geprüft werden.

Eine andere Möglichkeit eines Nachfolgeexperimentes besteht in der Validierung des *magic experiment*. So könnten etwa die abhängigen Variablen und der Aufgabenkontext variiert werden. Beispielsweise könnte den Versuchspersonen die chemische Manipulation offen dargelegt und eingangs des Experiments auch demonstriert werden. Die Realitätseinschätzung bzw. die Diskriminierungsfähigkeit würde dann nicht mehr in einen Zauberkontext gestellt, sondern z. B. mit der Frage „wie fest ist es getrickst?“ erfasst werden. Dadurch würde die Frage deutlicher auf die Diskrepanz zwischen erwarteter und beobachteter Intensität gelenkt. Es wäre interessant, festzustellen, ob diese Änderung bei den Kindern verschiedenen Alters zu anderen Einschätzungen führen würde.

Nahe liegend ist auch die Anwendung der Methode des *magic experiment* bei anderen Misskonzepten, deren Entwicklungsverlauf und Stabilität von Interesse ist. Es bestehen aufgrund der vorliegenden Daten grosse Hoffnungen, dass mit dem *magic experiment* eine einzigartige Methode gegeben ist, die Konflikthaftigkeit der teilweise sehr robusten Fehlvorstellungen bei jüngeren Kindern aufzudecken.

Weiter könnte die Methode der Präsentation widersprüchlicher Mischungsresultate bzw. verschiedener Rückmeldungen auch zur Überprüfung von differenziellen Lerneffekten bei der Mischung intensiver Grössen benutzt werden. Ein Paradigma, das auf den

Nachweis differenzieller Trainingseffekte abzielt, wurde von Baillargeon (1996, 2002) zur Untersuchung des physikalischen Wissenserwerbs von Säuglingen entwickelt. Die Grundidee dieser Methode besteht darin, dass zwei Gruppen von Kindern verschiedene Lernerfahrungen sammeln können: Einerseits in der regulären Trainingsbedingung durch Rückmeldung des korrekten Ereignisses, andererseits in der inversen Trainingsbedingung durch Rückmeldung eines manipulierten Ereignisses (nach Krist, Bach, Öndül & Huber, eingereicht). Ein differenzieller Trainingseffekt liegt dann vor, wenn die Kinder die reguläre, physikalisch korrekte Regel eher als die unkorrekte gelernt haben. Differenzielle Lerneffekte werden auf das Vorhandensein von relevantem Vorwissen zurückgeführt, das zu einer spezifischen Lernbereitschaft führt und mit dem aktuellen Wissenserwerb interagiert. Bei Krist, Bach, Öndül und Huber (eingereicht) z. B. zeigte sich ein signifikanter Effekt zu Gunsten der regulären Trainingsbedingung und zu Ungunsten der inversen Trainingsbedingung bei Trainingsexperimenten mit der Balkenwaage. Damit wurde die Annahme gestützt, dass aufgrund eines intuitiven, nicht explizierten Vorwissens über das Prinzip die Kinder leichter eine physikalisch korrekte als die hierzu inverse Regel erwerben konnten.

4.2 Kausalitätsverständnis gezauberter Ereignisse

Aus dieser Sicht standen nicht mehr primär die Erfassungen der Mischungsregeln bei der Mischung intensiver Grössen im Vordergrund, sondern die Erfassung des kindlichen Verständnisses über die Machbarkeit der verschiedenartigen Mischungsergebnisse in Abhängigkeit von unterschiedlich befähigten Akteuren. Das vorliegende Kapitel bezieht sich auf die erzielten Ergebnisse in Phase 3.

Grundlegende allgemeine Fragestellungen waren: Was wissen die Kinder über die Entstehungsbedingungen physikalisch möglicher und unmöglicher Farbmischungen? Wissen sie, dass die unmöglichen Mischungen anderen physikalischen Gesetzmässigkeiten folgen als die möglichen Mischungen? Was braucht es, um die gezauberten Ereignisse reproduzieren zu können? Gehören die präsentierten Ereignisse in den Bereich alltäglicher Erfahrung oder gehören sie in einen anderen Realitätsbereich, der den Kindern und möglicherweise auch anderen „normalen“ Personen nicht zugänglich ist?

Um diesen noch allgemeinen Fragestellungen auf die Spur zu kommen, wurden experimentell verschiedene Kausalbezüge zu den Entstehungsbedingungen zweier Mischungsarten hergestellt. Diese wurden unterschiedlich attribuierten Personen zugewiesen.

Die Attribute werden wie folgt umrissen:

- Versuchsperson selbst: Hier kommt die eigene Erfahrung, eigenes Wissen und Selbsteinschätzung der eigenen Fähigkeiten zum Tragen.
- Eltern der Versuchsperson: Bei den Zuschreibungen zu den Eltern wurde unterstellt, dass sie mehr Erfahrung und wahrscheinlich auch mehr Wissen als das Kind selbst besitzen. Der Erfahrungsbereich mit den elterlichen Fähigkeiten kann als gross eingeschätzt werden.
- Lehrperson: Hier wurde angenommen, dass diese nicht nur über mehr Erfahrung, sondern auch über mehr Wissen verfügen als die Versuchsperson. Der Erfahrungsbereich mit den Fähigkeiten der Lehrperson ist geringer als derjenige mit

den Eltern. In der Regel suggerieren Lehrpersonen ihren Schülern nicht, dass sie übernatürliche Fähigkeiten besitzen würden.

- **Zauberer:** Die wesentliche Eigenschaft der Figur des Zauberers besteht darin, dass dieser die Naturgesetze scheinbar durchbrechen kann. Er verfügt im spezifischen Bereich des gezeigten Ereignisses immer über mehr Wissen, Fertigkeiten oder Fähigkeiten als die Versuchsperson. Falls dies nicht der Fall ist, kann der Trick durchschaut werden und der Zauberer scheitert in seiner Rolle. Auch typisch für viele Zauberer – dies gilt aber nicht für alle – ist es, dass sie ihr Publikum gerne in der Illusion zurücklassen, dass sie nicht nur über besondere Fertigkeiten und Kenntnisse, sondern möglicherweise auch über übernatürliche Kräfte verfügen.

Auf der Basis dieser Attribuierungen wird davon ausgegangen, dass die Befähigung zur Erzeugung eines gezauberten Ereignisses in der obigen Reihenfolge immer mehr zunimmt. Die Zuschreibung von Befähigungen hängt erheblich von der Einsicht in den Entstehungsmechanismus des Ereignisses und der Vertrautheit mit dem Ereignis ab.

Es wird im Wesentlichen davon ausgegangen, dass sich bei uneinheitlichen Machbarkeitszuschreibungen sowohl das Ausmass der Einsicht in den Entstehungsmechanismus als auch die Art des Realitätsverständnisses der Versuchsperson widerspiegelt. Folgende Zusammenhänge werden postuliert:

- **Die Versuchsperson selbst ist ebenfalls zum Ereignis befähigt:** Es besteht hohe Einsicht in die Entstehungsbedingungen oder zumindest eine grosse Vertrautheit mit dem Ereignis. Es wird davon ausgegangen, dass bei hohen eigenen Machbarkeitszuschreibungen auch die anderen Akteure als hoch befähigt eingeschätzt werden.
- **Die Lehrperson ist zum Ereignis befähigt:** Die Lehrperson verfügt über die Fähigkeit und das Wissen, das Ereignis produzieren zu können. Das Ereignis gehört in den alltäglichen Realitätsbereich.
- **Nur der Zauberer ist zum Ereignis befähigt:** Nur der Zauberer verfügt über das Wissen, die Fertigkeit oder die Fähigkeit das Ereignis zu reproduzieren. Dahinter könnten verschiedene Kausalbezüge stecken: Beruht die Entstehung des Ereignisses „nur“ auf spezifischem Wissen oder spezifischer Fertigkeit (denkbar wäre etwa das Beschaffen der chemischen Substanz) so befindet sich der hergestellte

kausale Zusammenhang immer noch im Rahmen der alltäglichen Realität, im Sinne Subbotskys wurde ein natürlicher physikalischer Kausalzusammenhang hergestellt (Definition siehe weiter unten). Das physikalisch unmögliche Ereignis kann aber auch als gezaubertes Ereignis im Sinne des magischen Denkens interpretiert worden sein. In diesem Fall beruht die Ursachenzuschreibung auf einer magischen Kausalitätsbeziehung, infrage kämen etwa übernatürliche Fähigkeiten oder symbolische Handlungen, die das Ereignis bewirkt hatten. Ist nur der Zauberer zum Ereignis befähigt, so kann daraus geschlossen werden, dass weder Einsicht in den Entstehungsmechanismus noch Vertrautheit mit dem Ereignis besteht. Auch wenn dabei ein natürlicher physikalischer Kausalzusammenhang nicht ausgeschlossen werden kann, so würde das gezauberte Ereignis doch in einem erweiterten Realitätsbereich gesehen, wo mehr möglich ist als im Normalfall.

Die Versuchspersonen mussten in der vorliegenden Untersuchung sowohl die physikalisch möglichen als auch die unmöglichen Mischungsergebnisse anhand von Machbarkeitszuschreibungen beurteilen. Aus den resultierenden Zuschreibungen der beiden Mischungsarten bei Kindern verschiedenen Alters sowie bei Erwachsenen konnten einige Hinweise auf das Kausalitätsverständnis der untersuchten Personen herausgelesen werden.

Die allgemeinste Frage, ob die möglichen von den unmöglichen Ereignissen auf diesen Stufen überhaupt unterschieden wurden, kann eindeutig bestätigt werden. Ab dem Alter von acht Jahren unterschieden die Kinder die beiden Mischungsarten, dies allerdings mehrheitlich nur im Bereich der normalen Personen.

Es zeichnete sich aber nicht nur eine Entwicklungsveränderung der Unterscheidung zwischen den beiden Mischungsarten ab, sondern es zeigte sich auch ein differenziertes Bild zwischen den Attribuierungen zu den verschiedenen Akteuren. Deshalb wurde versucht, das kindliche Kausalitäts- und Realitätsverständnis auf der Basis der Ergebnisse aus Phase 3 gemäß den untersuchten Altersstufen nachzuvollziehen:

- Sechsjährige: Die beiden Ereignisse wurden hinsichtlich eines Unterschiedes in den Entstehungsbedingungen nicht differenziert. Dennoch zeigten auch die Vorschulkinder eine Differenzierung in der Attribuierung verschieden befähigter Ak-

teure: Die Lehrperson wurde als leicht höher befähigt eingestuft und der Zauberer wurde als fast maximal befähigt eingestuft. Die Sechsjährigen hatten zwar keine Einsicht in die Entstehungsmechanismen, ihre deutliche Unterscheidung zwischen dem Zauberer und anderen Akteuren lässt aber auf ein Kausalitätsverständnis schliessen, das dem Zauberer eine besondere Wirkungskraft verleiht.

- Achtjährige: Ab dem Alter von acht Jahren machten sich die beiden unterschiedlichen Entstehungsbedingungen in den Urteilen der Kinder bemerkbar. Es kann wohl von einem gewissen Verständnis der beiden Mechanismen ausgegangen werden, von einer Einsicht kann aber noch nicht gesprochen werden, da das Vertrauen in die eigene Machbarkeit der physikalisch möglichen Mischungsart im Mittel immer noch recht gering ausgeprägt war. Vermutlich ist es angemessener von einer unterschiedlichen Vertrautheit mit den beiden Mischungsarten anstelle von Einsicht zu sprechen. Dieser Unterschied der Entstehungsbedingungen wirkte sich nur bei den „normalen“ Personen aus, nicht hingegen beim Zauberer. Der Zauberer wurde erneut fast maximal eingeschätzt und zwar bei beiden Mischungsarten. Dass der Lehrperson bei dieser Altersgruppe deutlich mehr Wirkungskraft zugeschrieben wurde als noch bei den Sechsjährigen, zeigte, dass sie die Wirkungsbereiche der beiden Akteure und Lehrperson als nicht mehr so weit auseinander liegend einschätzten. Ein rudimentäres Verständnis der Natur des gezauberten Ereignisses in Richtung natürlicher physikalischer Kausalität könnte ein Grund für diese Erhöhung sein.
- Zehnjährige: Der grösste Unterschied zu den Einschätzungen der Achtjährigen besteht in der Differenzierung beim Zauberer. Hier machten sich die beiden unterschiedlichen Entstehungsmechanismen bemerkbar. Der Zauberer soll das gezauberte Ereignis nicht mehr eindeutig bewirken können. Diese Einschätzung widerspricht dem allgemein gültigen Verständnis eines Zauberers. Wer ausser dem Zauberer sollte dann dazu befähigt sein? Ist dieses Resultat Ausdruck eines veränderten Bildes des Zauberers? Werden Zauberer in diesem in diesem Alter als Individuen mit je eigenem Wissen und eigenen Fähigkeiten betrachtet?
- Zwölfjährige: Die differenzierte Einschätzung des Zauberers hielt auch bei den Zwölfjährigen an: Der Zauberer war wohl zu den physikalisch möglichen Ereignissen maximal befähigt, nicht mehr aber so deutlich zu den physikalisch unmöglichen. Im Unterschied zu den Zehnjährigen kann bei den Sechstklässlern von einer vollständigen Einsicht in den Entstehungsmechanismus des physikalisch mög-

lichen Ereignisses gesprochen werden, da sowohl sie selber als auch die Lehrperson als hoch befähigt eingeschätzt wurden. Demgegenüber war das Verständnis für die physikalisch unmögliche Mischungsart in der Entwicklungsspanne von acht bis zwölf Jahren nicht grösser geworden.

- Erwachsene: Bei den Erwachsenen hingegen wurde dem Zauberer wieder bei beiden Mischungsarten sehr hohe Wirkungskraft zugeschrieben. Erst bei den erwachsenen Studierenden kann von einer Einsicht in den Entstehungsmechanismus der physikalisch unkorrekten Mischungsarten gesprochen werden. Da aber auch sie sich und ihren Eltern die physikalisch unkorrekten Ereignisse weniger zutrauten als die korrekten, scheint es sich dabei, um einen Mechanismus zu handeln, der nicht bei allen zum Allgemeinwissen gehört.

Die erhaltenen Kausalattributionen zu verschiedenen befähigten Personen liessen interessante Entwicklungsverläufe erkennen und zeigten, dass Rückschlüsse auf das Kausalitätsverständnis der Kinder gerechtfertigt zu sein scheinen.

Die wohl interessantesten Ergebnisse bestehen in den hohen kindlichen Zuschreibungen zum Zauberer selbst, deren hohen Differenzen zu den Zuschreibungen der anderen Akteure sowie die Veränderungen der Zuschreibungen zum Zauberer bei den älteren Kindern.

Das Ergebnis, dass dem Zauberer besondere Wirkungskraft zugeschrieben wurde, war zu erwarten gewesen. Bereits durch die Untersuchung von Chandler und Lalonde (1994) mit neun- bis dreizehnjährigen Kindern war gezeigt worden, dass der Zauberer ein unmögliches Ereignis maximal plausibilisiert und ihm auch von älteren Kindern die Fähigkeit zugeschrieben wurde, ein scheinbar unmögliches Ereignis produzieren zu können. Chandler und Lalonde (1994) hatten dieses Ergebnis ins Licht der Magie als eine Art Restkausalitätskategorie gerückt. Magische Erklärungen würden immer dann herangezogen, wenn keine anderen Erklärungen vorhanden seien. Erklärungen oder Zuschreibungen, die zur Kategorie der Magie gezählt werden, sind aber nicht mit magischem Denken gleichzusetzen. Auch die Überzeugung, dass es sich beim unmöglichen Ereignis um eine Täuschung durch unbemerkte und unbekannte Manipulation handelte, führte zu derartigen Erklärungen oder Zuschreibungen. Die Vermutung, dass bei den Kindern – zumindest bis zum Alter von acht Jahren – allerdings ein erheblicher Anteil

der hohen Zuschreibungen zum Zauberer auf magisches Denken zurückzuführen ist, wird durch verschiedene Argumente bestärkt: Erstens gibt es einige Studien bei älteren Kindern und Erwachsenen, die magisches Denken nicht nur bei Vorschulkindern nachweisen konnten (v. a. die Arbeiten von Subbotsky, siehe weiter unten). Zweitens berichteten verschiedene Zauberer, die vor Kindern zaubern, dass Kinder bis zum Alter von acht Jahren mehrheitlich davon überzeugt seien, dass sie richtig zaubern könnten. Drittens weisen die Erfassungen der Erklärungen zur Entstehung der gezauberten Mischungsergebnisse der vorliegenden Untersuchung darauf hin, dass mit einem erheblichen Anteil an magischem Denken zu rechnen ist. Methodisch kann dieser Erfassung wohl angelastet werden, dass die zweite Frage im Anschluss an die offene erste Frage etwas suggestiv war. Hier kann aber entgegnet werden, dass auch bei derjenigen Versuchsleiterin, welche nur die offene erste nicht aber die zweite Frage gestellt hatte, ein Anteil an magischen Erklärungen erhoben wurde.

Auffällig waren auch die hohen Differenzen bei den jüngeren Kindern zwischen der hoch attribuierten Machbarkeit durch den Zauberer und den viel tieferen Werten der anderen „normalen“ Akteure. Es lässt sich vermuten, dass der Zauberer einer Kausalitätskategorie angehört, die sich von derjenigen normaler Personen und deren Wirkungsbereich im normalen Alltag unterscheidet. Gleichzeitig liegt es nahe, dass dieser Unterscheidung auch eine nicht vorhandene Einsicht in den Entstehungsmechanismus der unmöglichen Ereignisse zu Grunde liegt. Die Magie als eine Art Restkausalitätskategorie scheint auch hier angewendet worden zu sein. Weiter ist denkbar, dass magische Kausalbezüge im Sinne des magischen Denkens vorgenommen und dabei die Wirkungsbereiche von normalen Personen und von Zauberern sehr deutlich unterscheiden wurden.

Das magische Denken bzw. ein schwindendes magisches Denken im spezifischen Aufgabenkontext dieser Untersuchung könnte auch der Grund für das veränderte Bild des Zauberers bei den älteren Kindern gewesen sein. Die erhöhte Skepsis dieser Kinder gegenüber den Fähigkeiten des Zauberers könnte Ausdruck davon sein, dass der Zauberer seiner Allmachtsposition enthoben worden war, die er in jüngeren Jahren noch innehatte.

Die Ergebnisse lassen den Schluss zu, dass mit den Machbarkeitsattributionen weitere Erkenntnisse über die Art und die Entwicklung des Kausalitätsverständnisses gewonnen werden können. Um die Ergebnisse in einen weiteren Theoriebereich einzubetten und den Ausblick auf weitere Forschung zu fundieren, wird an dieser Stelle ein kleiner Abriss über den heutigen Stand der Forschung zur Entwicklung des Kausalitätsverständnisses mit besonderer Berücksichtigung auf verschiedene Kausalitätskategorien gegeben.

Heutige Ansicht zur Entwicklung des Kausalitätsverständnisses

Das Verständnis der Entwicklung des kausalen Denkens war lange Zeit durch das Werk Piagets (1975) geprägt.

Piagets Vorstellungen gelten heute aber in wesentlichen Punkten als überholt. Die bedeutendste Änderung besteht in einer grundlegenden Trennung der Entwicklung des kausalen Denkens von verschiedenen Inhaltsbereichen. Im Wesentlichen wird das kausale Denken über psychische Zustände und Erlebnisse (naive Psychologie) vom kausalen Denken über physikalische Objekte und Ereignisse (naive Physik) unterschieden. Diese beiden Bereiche würden sich zwar gewöhnlich parallel zueinander entwickeln, dennoch würden sich die kindlichen Theorien in den beiden Bereichen fundamental voneinander unterscheiden. Jedem Bereich komme die Bedeutung eines eigenen grundlegenden Wissensbereiches zu. Die Kausalität wird dabei als elementarer Entwicklungsmechanismus gesehen, der die Entwicklung in beiden Wissensbereichen beeinflusst, indem das kausale Denken z. B. die Aneignung weiterer Konzepte anstosse, forme und in bestimmte Bahnen lenke (Goswami, 2001).

Umgekehrt wird auch dem bereichsspezifischen Wissen bei der Kausalitätsentwicklung grosse Bedeutung zugeschrieben (vgl. z. B. Wellman & S. Gelman, 1992). Falsche kausale Bezüge können auf einem Mangel an Wissen basieren, wie das bei jüngeren Kindern sehr häufig der Fall ist (Sodian, 1998). Diesbezügliche Hinweise kamen z. B. von Rochel Gelman (1978), die zeigen konnte, dass Kinder mit höherer Vertrautheit mit dem Kontext einer Ursachen-Wirkungsaufgabe auch ein differenzierteres Kausalverständnis in diesem Bereich hatten als andere Kinder. Bullock, R. Gelman und Baillargeon (1982) unterscheiden folgende zwei Quellen des Wissens beim Erwerb von Kausalbezügen: Erstens das Erlernen von Kausaltheorien im Sozialisationsprozess, indem die

Eltern im täglichen Umgang mit ihren Kindern immer wieder auf Ursachen-Wirkungsprinzipien hinweisen. Zweitens die Erfahrung durch direktes, wiederholtes Beobachten der Umwelt. In beiden Fällen kann davon ausgegangen werden, dass das Wissen mit zunehmendem Alter immer umfassender wird und die Kausalbeziehungen auch adäquater wahrgenommen werden.

Spezifisches Wissen über die Entstehungsbedingungen eines Ereignisses beeinflusst die kausale Ursachenattribution des Ereignisses demnach entscheidend. Dies wurde auch in der vorliegenden Untersuchung deutlich: Magische kausale Zusammenhänge im Sinne des magischen Denkens wurden besonders bei denjenigen Kindern vermutet, die nicht über spezifisches Wissen verfügen. Dass die Kategorie „Magie“ von vielen Forschern als eine Restkategorie betrachtet wird, die dann zum Einsatz kommt, wenn auf keine anderen plausiblen Erklärungen zurückgegriffen werden kann, wurde schon mehrmals erwähnt. Dennoch greift diese Erklärung nach heutiger Ansicht in einigen Fällen zu kurz.

Sodian (1998) illustriert am Beispiel des kindlichen Animismus, warum die mangelnde Akkumulation von isolierten Wissensbeständen nicht alleine zur Erklärung fehlerhafter Kausalbezüge ausreicht, sondern auch tief greifende – ob bereichsspezifische und metakognitive ist noch unklar – Restrukturierungsprozesse in Betracht gezogen werden sollten. Animistische Erklärungen (Objekten werden psychische Zustände zugeschrieben, die handlungsbestimmend sind) sind nicht einfach nur Ausdruck eines mangelnden Wissens, sondern sie bedienen sich eines Begriffsystems, das aus einem anderen Bereich stammt – dem Bereich mentaler Zustände – und von Erwachsenen nicht zur Erklärung desselben Problems herangezogen wird. Damit sich aber dieses alternative Denken der Kinder verändert, braucht es fundamentale Veränderungen wie etwa eine Modifikation des Gesamtsystems, bei der komplexe begriffliche Systeme verändert werden. Dies würde erklären, weshalb die falschen kindlichen Kausalbezüge so resistent gegenüber Instruktion oder Wissensvermittlung sind.

In einer pädagogisch orientierten empirischen Studie mit Vorschulkindern von Sondermeyer (1978) wurde gezeigt, dass eine gezielte Aufklärung der Kinder darüber, dass der Zauberer nicht richtig zaubert, keinen Einfluss auf die magischen Vorstellungen der Kinder hatte und sie auch nach der Aufklärung davon überzeugt waren, der Zauberer

habe in der eben gesehenen Zaubervorstellung richtig gezaubert. Hier stellt sich die Frage, ob Kinder mit noch unvollständigen metakognitiven Fähigkeiten überhaupt die Darbietungen des Zauberers als Illusion erkennen können. Wie soll das Kind nicht an die Realität des Gezeigten glauben, wenn es gar nicht verstehen kann, dass das Gesehene nur Schein ist und nicht der Wirklichkeit entspricht? Die Fähigkeit zur Unterscheidung von Schein und Wirklichkeit kann zwar schon bei Kindern im Alter von etwa vier Jahren erstmals festgestellt werden (z. B. Flavell, 1983; Flavell, Green & Flavell, 1986; Taylor & Flavell, 1984). Wie schnell diese Fähigkeit aber im Umgang mit der komplexen sozialen Umwelt eingesetzt werden kann, ist noch kaum bekannt (vgl. Flavell, 1989; Russell, 1985, 1988). Auch Sodian weist darauf hin, dass wir heute noch sehr wenig über Veränderungen der kindlichen Theory of Mind – und die Unterscheidung von Schein und Wirklichkeit ist eine zentrale Fähigkeit im Rahmen der Theory of Mind – nach Ende der Vorschulzeit wissen. Sie stellt selbst die Hypothese auf, dass Kinder zunehmend Einsicht in den konstruktiven Charakter geistiger Aktivität gewinnen und diese Entwicklung im Jugendalter dazu führt, das Realitätsverständnis selbst als Konstruktion zu betrachten und infrage zu stellen.

Aus dieser Sicht betrachtet, scheint es nicht abwegig, das magische Denken – allerdings in Abhängigkeit von der spezifischen Situation – auch noch im Schulalter zu erwarten. Möglicherweise kommt diesem „Denkfehler“, auch wenn er im normalen Alltag kaum zu Tage tritt, sogar eine wichtige Bedeutung in der Erforschung der Entwicklung des kausalen Denkens zu.

Aus dieser Sicht kann die magische Kausalität als missinterpretierter „Denkfehler“ verstanden werden, über dessen Hintergrund aber noch wenig bekannt ist. Interessant ist nun, diese Hypothesen mit einer anderen theoretischen Konzeption in Verbindung zu setzen, die ausgehend von magischer Kausalität entwickelt wurde und eine ganz andere Entstehungsgeschichte hat.

Theoretische Konzeption zur Entwicklung des Kausalitätsverständnisses unter Berücksichtigung der magischen Kausalität

Es liegen seit den neunziger Jahren zahlreiche Studien vor, die immer wieder auf ein magisches Denken bei Vorschulkindern, aber auch bei älteren Kindern und Erwachsenen, schließen lassen. Die jeweiligen theoretischen Schlussfolgerungen reichten aber kaum über die Debatte, ob es nun ein magisches Denken gebe oder nicht, hinaus. Davon

heben sich allerdings die neueren Arbeiten von Subbotsky (u. a. 2001) ab. Seine theoretische Konzeption zur Entwicklung des Kausalitätsverständnisses unter Berücksichtigung der magischen Kausalität stützt sich auf mehrere eigene Studien ab (z. B. Subbotsky 1992, 1994, 1997a, 1997b, 2000b, 2000c, 2001, 2004; Subbotsky und Trommsdorff, 1992).

Zentral an Subbotskys Ansatz ist, dass die Kausalitätsentwicklung nicht mehr wie bei Piaget (1930, 1986) und vielen anderen Autoren (für eine Übersicht s. White, 1995; nach Subbotsky, 2000a) einzig und allein in ein Denken von natürlicher physikalischer Kausalität (Definition des Begriffes siehe unten) mündet, sondern daneben noch andere, allerdings weniger dominante Kausalitätskategorien (wie z. B. die magische Kausalität oder der Animismus) erhalten bleiben. Subbotsky verwirft das *Replacement*-Modell und postuliert stattdessen ein Koexistenz-Modell verschiedener Kausalitätsarten.

Subbotsky (2000a) unterscheidet verschiedene Arten von kausalen Bezügen, die er zwei übergeordneten Kategorien zuordnet. Zentral ist demnach die Unterscheidung zwischen physikalischer und psychologischer Kausalität.

Psychologische Kausalität tritt auf, wenn die Ursache (A) für das Ereignis (B) ein Subjekt ist. Zur Kategorie der psychologischen Kausalität gehört die psychophysische, magische, artifizielle und soziale Kausalität. Psychophysische Kausalität ist dann der Fall, wenn ein Gedanke des Subjekts eine körperliche Bewegung verursacht. Im Falle der magischen Kausalität bewirkt der Gedanke oder eine symbolische Handlung eines Subjektes ein Ereignis in der physikalischen Welt. Artifizielle Kausalität ist gegeben, wenn die physikalische Handlung eines Subjektes direkt oder durch gedankliche Assoziation der Grund für eine Veränderung in der Welt der Natur darstellt (z. B. Windmühlen werden als Verursacher des Windes verstanden). Die soziale Kausalität schliesslich tritt ein, wenn die Handlung eines Subjektes (physisch oder verbal) ein Verhalten eines anderen Subjektes verursacht.

Physikalische Kausalität hingegen bezieht sich auf die kausale Verbindung zweier physikalischer Ereignisse. Hierzu nennt Subbotsky zwei Kausalitätsarten: Zum einen spricht er von spontaner physikalischer Kausalität, welche auftritt, wenn das Ereignis A die „causa finalis“ des Effektes B ist (z. B. wenn sich die Aktivitätsquelle im Objekt /

Subjekt selbst befindet, wie es normalerweise nur bei lebendigen Objekten / Subjekten der Fall ist und keine andere Ursache infrage kommt). Der Animismus, bei dem unbelebten Objekten Leben oder Bewusstsein zugeschrieben wird, kann als fehlerhafter Kausalbezug im Sinne der spontanen physikalischen Kausalität verstanden werden. Die zweite Kausalitätsart im Rahmen der physikalischen Kausalität ist die natürliche physikalische Kausalität, welche immer zu einer unbegrenzten Verkettung von Ursache-Wirkungsbeziehungen gehört.

Auch für Subbotsky spielt die Ausbildung und Entwicklung der Theory of Mind in der Kausalitätsentwicklung eine entscheidende Rolle. Sie eröffnet im kindlichen kausalen Denken eine neue – immaterielle – Dimension, indem das Denken über Ursachen des Verhaltens von Subjekten, die z. B. eigene Absichten haben, möglich wird. So ist die Ursache des Verhaltens anderer Menschen nicht in einem Naturgesetz zu suchen, sondern das Verhalten anderer Menschen hängt nunmehr von ihrem situationellen Wissen ab. Die Fähigkeit, die subjektiven psychischen Zustände anderer Menschen als verhaltenssteuernd zu betrachten, ist für die Entwicklung psychologischer Kausalität zentral.

Subbotsky verwehrt sich aufgrund seiner empirischen Basis gegen das vollständige Verschwinden des animistischen und magischen Denkens im Verlaufe der Kindheit, wie es das *Replacement*-Modell postuliert hatte. Er versteht diese Kausalitätsbezüge nicht als „Denkfehler“, die auf einen Mangel an Wissen zurückzuführen wären, sondern als Ausdruck einer spezifischen Eigenart des Denkens, die tief im menschlichen Denken verwurzelt sei. Sie treten alternativ zum wissenschaftlichen Denken auf, gerade auch in Situationen, in denen keine wissenschaftliche Erklärung existiert oder wenn diese sehr kompliziert ist. Subbotsky konnte durch seine Untersuchungen weitere Situationsmerkmale ausfindig machen: Ein Ereignis darf möglichst nicht technologisch erklärt werden können, weshalb die technologische Erklärbarkeit experimentell minimal gehalten werden sollte. Weiter sollte der Ausdruck oder die Äusserung von magischem Denken keine negativen sozialen Sanktionen zur Folge haben (z. B. Subbotsky, 2000a; Woolley & Phelps, 1994). Ferner manipulierte Subbotsky in einer seiner neuesten Untersuchungen (2001) die Höhe eines allfälligen Schadens, der sich für die Versuchsperson bei magischem Kausalbezug ergeben würde. Es zeigte sich, dass bei hohem erwarteten Schaden, auch die erwachsenen Versuchsteilnehmer zu deutlich leichtgläubigem Verhalten neigten. Schliesslich wiesen nebst Subbotsky (z. B. 1985, 1994) auch andere Autoren darauf

hin, dass ein magischer Kontext magische Kausalitätsbezüge fördere (Rosengren & Hickling, 1994). Eine systematische Analyse der bestehenden Literatur würde an dieser Stelle gewiss noch weitere Situationsmerkmale zur Evokation von magischem Denken aufdecken können.

Zusammenfassend lässt sich schlussfolgern, dass magisches Denken und Animismus immer wieder experimentell und in starker Abhängigkeit von der Aufgabensituation nachgewiesen werden konnten. Interessant scheint demnach für die Zukunft weniger die Beantwortung der Frage, ob es magisches Denken gibt, sondern vielmehr die Identifizierung der auslösenden Bedingungen und der Arten von Kausalitätskategorien.

Bei der Erforschung dieser Kategorie und Art von Denken könnte die entwicklungspsychologische Perspektive sehr hilfreich sein, da nicht nur in der vorliegenden Untersuchung sondern eigentlich in allen Studien zum magischen Denken mit Kindern Altersunterschiede und Entwicklungsveränderungen gefunden wurden.

Mögliche entwicklungsbedingte Veränderungen im kausalen Verständnis der magischen Kategorie könnten durchaus auch in den generell wichtigen kognitiven Veränderungen wie etwa der Entwicklung der Theory of Mind bzw. Fähigkeit zur Metakognition (ein Kind kann die Darbietung des Zauberers ohne ausgereifte Theory of Mind nicht als Illusion erkennen) oder auch dem Wissenserwerb (ohne Kenntnis der zu Grunde liegenden Manipulation bleibt die Interpretation eines unmöglichen Ereignisses spekulativ) liegen.

Fruchtbar wäre vermutlich auch ein Vergleich mit den Befunden zum magischen Denken aus den Forschungsbereichen Psychopathologie (z. B. Eckblad & Chapman, 1983; Thalbourne, 1994; Vyse, 1997) und der Neurowissenschaften (z. B. Brugger & Taylor, 2003; Pizzagalli, Lehmann, Gianotti, König, Tanaka, Wackermann & Brugger, 2000), die nach neurobiologischen Korrelaten Ausschau hält.

Ganz am Schluss soll noch der Frage nachgegangen werden, welches Forschungspotenzial die Machbarkeitszuschreibungen zu verschiedenen Akteuren, wie sie in der vorliegenden Untersuchung verwendet wurden, noch haben könnten. Nahe liegend scheint es, nach dem Grund für die Veränderung der Einstellung zum Zauberer bei den älteren Kindern zu forschen. Denkbar wäre es, direkt nach magischem Denken über Machbarkeitszuschreibungen zu suchen. Um einen eindeutigen Zusammenhang zwischen einer hohen Machbarkeitszuschreibung zum Zauberer und magischer Kausalität nachweisen

zu können, müssten die experimentellen Bedingungen allerdings verändert werden. Bei vorliegendem Experiment waren die Machbarkeitseinschätzungen – aus dieser Sicht betrachtet – mit dem Farbmischungsverständnis der Versuchsperson und der Einsicht in den Entstehungsmechanismus des unmöglichen Ereignisses konfundiert. Würde man beides eliminieren, indem anstelle von Farbmischungen nur Mischungen von farblosen Flüssigkeiten gezeigt würden und die gefüllten Gläser mit dem unsichtbaren Zusatz der Versuchsperson zur Verfügung stehen würden, wäre bei einer ausschliesslichen Machbarkeitszuschreibung zum Zauberer eigentlich nur noch eine magische Kausalitätsverknüpfung denkbar. Bei vorhandenen Flüssigkeiten bzw. vorhandenem Trickmaterial und der Gewissheit, dass die Farbänderung an den Flüssigkeiten gelegen hatte, sollte sich die Versuchsperson ohne magisches Denken auch selbst zur Reproduktion des Gezeigten befähigt fühlen. Wenn sie sich selbst und andere „normale“ Personen im Gegensatz zum Zauberer aber nicht als dazu befähigt sehen, kann es nur an den besonderen Fähigkeiten des Zauberers liegen. Möglich wäre allerdings, dass der Zauberer – im Vergleich zur Präsentation in der vorliegenden Untersuchung – seine „Hilfsmittel“ wie etwa eine magische Atmosphäre oder einen Zauberspruch zur Hilfe nehmen müsste, um eine deutlichere magische Kausalität zu suggerieren.

5 Literaturverzeichnis

- Acredolo, C. & Horobin, K. (1987). The child's relational reasoning and tendency toward premature closure. *Developmental Psychology*, 23, 13-21.
- Acredolo, C. & O'Connor, J. (1991). On the difficulty of detecting cognitive uncertainty. *Human Development*, 34(4), 204-223.
- Aebli, H. (1984). Kognitive Entwicklung: Was entwickelt sich, und bei welchen Anlässen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 16, 102-118.
- Ahl, V. A., Moore, C. F. & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development*, 7(1), 81-108.
- Anderson, N. H. (1980). Information integration theory in developmental psychology. In F. Wilkening, J. Becker & T. Trabasso (Hrsg.), *Information integration by children* (S. 1-45). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.
- Anderson, N. H. (1987). Function knowledge: Comment on Reed and Evans. *Journal of Experimental Psychology: General*, 116(3), 297-299.
- Anderson, N. H. (1996). *A functional theory of cognition*. Mahwah, NJ: Erlbaum.
- Anderson, N. H. & Cuneo, D. O. (1978a). The height + width rule in children's judgments of quantity. *Journal of Experimental Psychology: General*, 107(4), 335-378.
- Anderson, N. H. & Cuneo, D. O. (1978b). The height + width rule seems solid: Reply to Bogartz. *Journal of Experimental Psychology: General*, 107(4), 388-392.
- Anderson, N. H. & Wilkening, F. (1991). Adaptive thinking in intuitive physics. In N. H. Anderson (Hrsg.), *Contributions to information integration theory* (Bd. 3, S. 1-42). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.

- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants. *Cognition*, 23(1), 21-41.
- Baillargeon, R. (1987). Object permanence in 3 1/2- and 4 1/2-month-old infants. *Developmental Psychology*, 23(5), 655-664.
- Baillargeon, R. (1993). The object concept revisited: New direction in the investigation of infants' physical knowledge. In C. Granrud (Hrsg.), *Visual perception and cognition in infancy. Carnegie Mellon symposia on cognition* (S. 265-315). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.
- Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. K. Rovee-Collier & L. P. Lipsitt (Hrsg.), *Advances in infancy research* (Bd. 9). Norwood, NJ: Ablex.
- Baillargeon, R. (1996). Infants' understanding of the physical world. In M. Sabourin, F. Craik & M. Robert (Hrsg.), *Advances in psychological science* (Bd. 2, S. 503-529). London: Psychology Press.
- Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Hrsg.), *Handbook of childhood cognitive development* (S. 47-83). Oxford: Blackwell.
- Baillargeon, R., Spelke, E. S. & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20(3), 191-208.
- Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Hrsg.), *Concepts and conceptual development: Ecological and intellectual factors in categorization. Emory symposia in cognition*, 1 (S. 101-140). New York, NY, US: Cambridge University Press.
- Barsalou, L. W. (1993). Challenging assumptions about concepts. *Cognitive Development*, 8(2), 169-180.
- Beilin, H. (1983). The new functionalism and Piaget's program. In E. K. Skolnick (Hrsg.), *New trends in conceptual representation: Challenges to Piaget's theory?* (S. 3-40). Hillsdale; NJ: Erlbaum.
- Beilin, H. (1989). Piagetian Theory. In R. Vasta (Hrsg.), *Annals of child development: Vol. 6, Six theories of child development: Revised formulations and current issues* (S. 85-131). Greenwich, CT: JAI Press.

- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Brosius, F. (2002). *SPSS 11*. Bonn: MITP.
- Brugger, P. & Taylor, K. I. (2003). ESP: Extrasensory perception or effect of subjective probability? In J. E. Alcock (Hrsg.), *Psi wars: Getting to grips with the paranormal*. (S. 221-246). Charlottesville, VA: Imprint Academic.
- Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist*, 19, 1-15.
- Bruner, J. S., Olver, R. R. & Greenfield, P. M. (1971). *Studien zur kognitiven Entwicklung*. Stuttgart: Klett.
- Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkeley: University of California Press.
- Bühl, A. & Zöfel, P. (2002). *SPSS 11. Einführung in die moderne Datenanalyse unter Windows*. München: Pearson Studium.
- Bullock, M., Gelman, R. & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Hrsg.), *The developmental psychology of time* (S. 209-254). New York: Academic Press.
- Butzin, C. A. & Anderson, N. H. (1973). Functional measurement of children's judgments. *Child Development*, 44(3), 529-537.
- Byrnes, J. P. & Beilin, H. (1991). The cognitive basis of uncertainty. *Human Development*, 34(4), 189-203.
- Carey, S. (1986). Cognitive science and science education. *American Psychologist*, 41(10), 1123-1130.
- Case, R. (1985). *Intellectual Development: Birth to Adulthood*. Orlando, FL: Academic Press.
- Chandler, M. J. & Lalonde, C. E. (1994). Surprising, magical and miraculous turns of events: Children's reactions to violations of their early theories of mind and matter. *British Journal of Developmental Psychology*, 12(1), 83-95.
- Chapman, M. (1988). *Constructive evolution: Origins and development of Piaget's thought*. Cambridge University Press.

- Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics*, 50, 66-71.
- DeVries, R. (1969). Constancy of generic identity in the years three to six. *Monographs of the Society for Research in Child Development*, 34(3), 56-60.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10(2-3), 105-225.
- Dixon, J. A. & Moore, C. F. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology*, 32(2), 241-253.
- Eckblad, M. & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology*, 51, 215-225.
- Elkind, D. & Schoenfeld, E. (1972). Identity and equivalence conservation at two age levels. *Developmental Psychology*, 6(3), 529-533.
- Field, D. (1987). A review of preschool conservation training: An analysis of analyses. *Developmental Review*, 7(3), 210-251.
- Fischer, K. W. (1991). The failure of competence: How context contributes directly to skills. In R. Wozniak & K. W. Fischer (Hrsg.), *Specific environments: Thinking in contexts*. Hillsdale, NJ: Erlbaum.
- Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology*, 2(4), 421-453.
- Flavell, J. H. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, 15(1), 95-120.
- Flavell, J. H., Flavell, E. R. & Green, F. L. (1989). A transitional period in the development of the appearance-reality distinction. *International Journal of Behavioral Development*, 12(4), 509-526.
- Flavell, J. H., Green, F. L. & Flavell, E. R. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development*, 51(1), 1-68.
- Gelman, R. (1972). Logical capacity of very young children: Number invariance rules. *Child Development*, 43(1), 75-90.
- Gelman, R. (1978). Cognitive Development. *Annual Review of Psychology*, 29, 297-332.

- Gelman, R. & Baillargeon, R. (1983). A review of some Piagetian concepts. In J. H. Flavell & E. M. Markman (Hrsg.), *Handbook of child psychology: Vol. III. Cognitive development* (S. 167-230). New York: Wiley.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Goswami, U. (2001). *So denken Kinder*. Bern: Huber.
- Hall, V. C. & Kaye, D. V. (1978). The necessity of logical necessity. In L. S. Siegel & C. J. Brainerd (Hrsg.), *Alternatives to Piaget: Critical essays on the theory* (S. 153-167). New York: Academic Press.
- Halloun, I. B. & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, 53, 1056-1065.
- Halloun, I. B. & Hestenes, D. (1985b). The initial knowledge state of college physics students. *American Journal of Physics*, 53, 1043-1055.
- Hammond, K. R. (1982). *Unification of theory and research in judgment and decision making*. Boulder: University of Colorado, Center for Research on Judgment and Policy.
- Hammond, K. R., Hamm, R. M., Grassia, J. & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions, SMC-17*, 753-770.
- Harris, P. L. (1994). Unexpected, impossible and magical events: Children's reactions to causal violations. *British Journal of Developmental Psychology*, 12, 1-7.
- Harris, P. L., Brown, E., Marriot, C., Whittal, S. & Harmer, S. (1991). Monsters, ghosts and witches: Testing the limits of the fantasy-reality distinction in young children. *British Journal of Developmental Psychology*, 9, 105-123.
- Hauser, M. D. & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology*, 47(4), 367-401.
- Huang, I. (1930). Children's explanations of strange phenomena. *Psychologische Forschung*, 14, 63-183.
- Hülksen, C., Sodian, B. & Pickel, G. (2001). Die Unterscheidung von Schein und Sein im Verkleidungsspiel: Ein Problem der dualen Kodierung oder der Identitätser-

- haltung? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 33(3), 129-137.
- Jäger, S. (1996). *Entwicklungsprozesse der kognitiven Durchschnittsbildung*. Münster: Waxmann.
- Jäger, S. (1997). Entwicklung des Regelerwerbs bei der Integration intensiver Grössen. In A. Ruemmele, S. Pauen & G. Schwarzer (Hrsg.), *Kognitive Entwicklungspsychologie: Aktuelle Forschungsergebnisse* (S. 129-144): Pabst, Lengerich.
- Jäger, S. & Becker, M. (1997). Die Entwicklung der Durchschnittsbildung als adaptiver Denkprozess: Prädiktion und Rekonstruktion von Farbintensitätsmischungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 29(4), 291-300.
- Jäger, S. & Wilkening, F. (2001). Development of cognitive averaging: When light and light make dark. *Journal of Experimental Child Psychology*, 79(4), 323-345.
- Johnson, C. N. & Harris, P. L. (1994). Magic: Special but not excluded. *British Journal of Developmental Psychology*, 12(1), 35-51.
- Kahneman, D. & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, 11, 123-141.
- Kosslyn, S. M. (1978). The representational-development hypothesis. In P. A. Ornstein (Hrsg.), *Memory development in children* (Bd. 157-189). Hillsdale, NJ: Erlbaum.
- Krist, H., Bach, S., Öndül, S. & Huber, S. (eingereicht). Mikrogenetische Studien zum physikalischen Wissenserwerb von Kindern: Neue Trainingsexperimente mit der Balkenwaage. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*.
- Krist, H., Fieberg, E. L. & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 952-966.
- Krist, H. & Wilkening, F. (1991). Repräsentationale Entwicklung. *Sprache & Kognition*, 10(4), 181-195.
- Lamsfuss, S. (1994). *Misskonzepte und Analogien: Kindliche Vorstellungen über das Zusammenwirken von Kräften*. Heidelberg: Asanger.

- Lindsay, D. S. & Creedon, C. F. (1985). "Magic" revisited: Children's responses to apparent violations of conservation. *Journal of Experimental Child Psychology*, 40(2), 338-349.
- Marschark, M., Greenberg, N. A. & Clark, M. D. (1983). Dimensions of number invariance. *Bulletin of the Psychonomic Society*, 21(2), 108-110.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Hrsg.), *Mental models*. Hillsdale, NJ: Erlbaum.
- McEvoy, J. & O'Moore, A. M. (1991). Number conservation: A fair assessment of numerical understanding? *Irish Journal of Psychology*, 12(3), 325-337.
- Mermelstein, E. & Shulman, L. S. (1967). Lack of formal schooling and the acquisition of conservation. *Child Development*, 38(1), 39-52.
- Miller, S. A. (1971). Extinction of conservation: A methodological and theoretical analysis. *Merrill Palmer Quarterly*, 17, 319-334.
- Miller, S. A. (1973). Contradiction, surprise, and cognitive change: The effects of disconfirmation of belief on conservers and nonconservers. *Journal of Experimental Child Psychology*, 15(1), 47-62.
- Miller, S. A. (1976). Extinction of Piagetian concepts: An updating. *Merrill Palmer Quarterly*, 22(4), 257-281.
- Miller, S. A. (1986). Certainty and necessity in the understanding of Piagetian concepts. *Developmental Psychology*, 22(1), 3-18.
- Miller, S. A., Brownell, C. A. & Zukier, H. (1977). Cognitive certainty in children: Effects of concept, developmental level, and method of assessment. *Developmental Psychology*, 13(3), 236-245.
- Moore, C. F., Dixon, J. A. & Haines, B. A. (1991). Components of understanding in proportional reasoning: A fuzzy set representation of developmental progressions. *Child Development*, 62(3), 441-459.
- Novak, J. D. (1987). *Proceedings of the second international seminar on misconceptions and educational strategies in science and mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Phelps, K. E. & Woolley, J. D. (1994). The form and function of young children's magical beliefs. *Developmental Psychology*, 30(3), 385-394.

- Piaget, J. (1928). La causalité chez l'enfant. *British Journal of Psychology*, 18, 276-301.
- Piaget, J. (1929). *The child's conception of the world*. London: Kegan Paul.
- Piaget, J. (1930). *The child's conception of physical causality*. London: Kegan Paul.
- Piaget, J. (1957). Logique et équilibre dans les comportements du sujet. In L. Apostel, B. Mandelbrot & J. Piaget (Hrsg.), *Etudes d'epistemologie génétique: Vol 2. Logique et l'équilibre* (S. 27-117). Paris: Presse Universitaires de France.
- Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green, M. P. Ford & G. B. Flamer (Hrsg.), *Measurement and Piaget* (S. 1-11). New York: McGraw Hill.
- Piaget, J. (1975). *Der Aufbau der Wirklichkeit beim Kinde* (Bd. 2). Stuttgart: Klett.
- Piaget, J. (1980). *Experiments in contradiction*. Chicago: University of Chicago Press. (Original erschienen 1974)
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Chicago: University of Chicago Press. (Original erschienen 1975)
- Piaget, J. (1986). *The construction of reality in the child*. New York: Ballantine Books.
- Piaget, J. & Inhelder, B. (1977). *Von der Logik des Kindes zur Logik des Heranwachsenden*. Olten: Walter.
- Piaget, J. & Szeminska, A. (1975). *Die Entwicklung des Zahlbegriffs beim Kinde* (Bd. 3). Stuttgart: Klett-Cotta. (Original erschienen 1941)
- Pizzagalli, D., Lehmann, D., Gianotti, L., Koenig, T., Tanaka, H., Wackermann, J. & Brugger, P. (2000). Brain electric correlates of strong belief in paranormal phenomena: Intracerebral EEG source and regional Omega complexity analyses. *Psychiatry Research: Neuroimaging*, 100(3), 139-154.
- Reed, S. K. & Evans, A. C. (1987). Learning functional relations: A theoretical and instructional analysis. *Journal of Experimental Psychology: General*, 116(2), 106-118.
- Robert, M. (1989). Reduction of demand characteristics in the measurement of certainty during modeled conservation. *Journal of Experimental Child Psychology*, 47(3), 451-466.

- Rosengren, K. S. (1991). As time goes by: Children's early understanding of growth in animals. *Child Development*, 62(6), 1302-1320.
- Rosengren, K. S. & Hickling, A. K. (1994). Seeing is believing: Children's explanations of commonplace, magical, and extraordinary transformations. *Child Development*, 65(6), 1605-1626.
- Rosengren, K. S., Kalish, C. W., Hickling, A. K. & Gelman, S. A. (1994). Exploring the relation between preschool children's magical beliefs and causal thinking. *British Journal of Developmental Psychology*, 12(1), 69-82.
- Russell, J. & Haworth, H. M. (1988). Appearance versus reality in dyadic interaction: Evidence for a lingering phenomenism. *International Journal of Behavioral Development*, 11(2), 155-170.
- Russell, J. & Mitchell, P. (1985). Things are not always as they seem: The appearance/reality distinction and conservation. *Educational Psychology*, 5(3-4), 227-238.
- Santos, L. R. & Hauser, M. D. (2002). A non-human primate's understanding of solidity: Dissociations between seeing and acting. *Developmental Science*, 5(2), F1-F7.
- Santos, L. R., Miller, C. T. & Hauser, M. D. (2003). Representing tools: How two non-human primate species distinguish between the functionally relevant and irrelevant features of a tool. *Animal Cognition*, 6(4), 269-281.
- Schär, P. (2004). *Die Stabilität der Additionsregel bei der Entwicklung der kognitiven Durchschnittsbildung: Rekonstruktion von Farbintensitätsmischungen*. Unveröff. Lizenziatsarbeit, Universität Zürich, Psychologisches Institut, Abt. Allgemeine und Entwicklungspsychologie.
- Schmid, P. (1998). *Mythologisierung und Entmythologisierung okkulturer Praktiken*. Unveröff. Seminararbeit, Universität Zürich, Theologisches Seminar, Abt. Religionsgeschichte und Religionswissenschaft.
- Siegler, R. S. (1986). *Children's Thinking*. Englewood Cliffs, NJ: Prentice-Hall.
- Silverman, I. W., Rose, A. P. & Phillis, D. E. (1979). The "magic" paradigm revisited. *Journal of Experimental Child Psychology*, 28(1), 30-42.

- Singh, R., Sidana, U. R. & Saluja, S. K. (1978a). Integration theory applied to judgments of personal happiness by children. *Journal of Social Psychology*, 105(1), 27-31.
- Singh, R., Sidana, U. R. & Saluja, S. K. (1978b). Playgroup attractiveness studied with information integration theory. *Journal of Experimental Child Psychology*, 25(3), 429-436.
- Singh, R., Sidana, U. R. & Srivastava, P. (1978). Averaging processes in children's judgment of happiness. *Journal of Social Psychology*, 104(1), 123-132.
- Smedslund, J. (1961). The acquisition of conservation of substance and weight in children: III. Extinction of conservation of weight acquired "normally" and by means of empirical controls on a balance. *Scandinavian Journal of Psychology*, 2, 85-87.
- Sodian, B. (1998). Entwicklung bereichsspezifischen Wissens. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie: Ein Lehrbuch* (4. Aufl., S. 622-653). Weinheim: Psychologie Verlags Union.
- Sondermeyer, M. (1978). Zauberei für Kinder im Vorschulalter, Probleme der Desillusionierung. *Schriftenreihe zur Zauberkunst*, 8.
- Spiro, R. J., Feltovich, P. J., Coulsen, R. L. & Anderson, D. K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconceptions in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Hrsg.), *Similarity and analogical reasoning*. Cambridge: Cambridge University Press.
- Stavy, R. (1981). Teaching inverse functions via the concentration of salt water solution. *Archives de Psychologie*, 49(191), 267-287.
- Stavy, R. & Berkovitz, B. (1980). Cognitive conflict as a basis for teaching quantitative aspects of the concept of temperature. *Science Education*, 64, 679-692.
- Stavy, R., Strauss, S., Orpaz, N. & Carmi, G. (1982). U-shaped behavioral growth in ratio comparisons. In S. Strauss (Hrsg.), *U-shaped behavioral growth*. (S. 11-36). New York: Academic Press.
- Strauss, S. (1982a). Ancestral and descendant behaviors: The case of U-shaped behavioral growth. In T. G. Bever (Hrsg.), *Regressions in mental development: Basic*

- phenomena and theories* (S. 191-220). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Strauss, S. (1982b). Introduction. In S. Strauss (Hrsg.), *U-shaped behavioral growth* (S. 1-9). New York: Academic Press.
- Strauss, S. (Hrsg.). (1982c). *U-shaped behavioral growth*. New York: Academic Press.
- Strauss, S., Ankori, M., Orpaz, N. & Stavy, R. (1977). Schooling effects on the development of proportional reasoning. In Y. H. Poortinga (Hrsg.), *Basic problems in cross-cultural psychology* (S. 129-137). Amsterdam: Swets & Zeitlinger.
- Strauss, S. & Klein, R. (1985). The development of children's concepts of hardness. *The Journal of Genetic Psychology*, 146(4), 483-494.
- Strauss, S. & Stavy, R. (1982). U-shaped behavioral growth: Implications for theories of development. In W. W. Hartup (Hrsg.), *Review of child development research* (Bd. 6, S. 547-599). Chicago: University of Chicago Press.
- Subbotsky, E. (1985). Preschool children's perception of unusual phenomena. *Soviet Psychology*, 23, 91-114.
- Subbotsky, E. (1992). *Foundations of the mind: Children's understanding of reality*. Hertfordshire, England: Harvester Wheatsheaf.
- Subbotsky, E. (1994). Early rationality and magical thinking in preschoolers: Space and time. *British Journal of Developmental Psychology*, 12(1), 97-108.
- Subbotsky, E. (1997a). Explanations of unusual events: Phenomenalistic causal judgments in children and adults. *British Journal of Developmental Psychology*, 15(Pt 1), 13-36.
- Subbotsky, E. (1997b). Understanding the distinction between sensations and physical properties of objects by children and adults. *International Journal of Behavioral Development*, 20(2), 321-347.
- Subbotsky, E. (2000a). Causal reasoning and behaviour in children and adults in a technologically advanced society: Are we still prepared to believe in magic and animism? In P. Mitchell & K. J. Riggs (Hrsg.), *Children's reasoning and the mind* (S. 327-348). Hove, East Sussex: Psychology Press Ltd.
- Subbotsky, E. (2000b). Phenomenalistic perception and rational understanding in the mind of an individual: A fight for dominance. In K. S. Rosengren (Hrsg.), *Imag-*

- ining the impossible: Magical, scientific, and religious thinking in children* (S. 35-74). New York: Cambridge University Press.
- Subbotsky, E. (2000c). Phenomenalistic reality: The developmental perspective. *Developmental Review*, 20(3), 438-474.
- Subbotsky, E. (2001). Causal explanations of events by children and adults: Can alternative causal modes coexist in one mind? *British Journal of Developmental Psychology*, 19(Pt1), 23-45.
- Subbotsky, E. (2004). Magical thinking in judgments of causation: Can anomalous phenomena affect ontological causal beliefs in children and adults? *British Journal of Developmental Psychology*, 22, 123-152.
- Subbotsky, E. & Trommsdorff, G. (1992). Object permanence in adults: a cross-cultural perspective. *Psychologische Beiträge*, 34(1-2), 62-79.
- Taylor, M. & Flavell, J. H. (1984). Seeing and believing: Children's understanding of the distinction between appearance and reality. *Child Development*, 55(5), 1710-1720.
- Thalbourne, M. (1994). Belief in the paranormal and its relationships to schizophrenia-relevant measures: A confirmatory study. *British Journal of Clinical Psychology*, 33, 78-80.
- Vyse, S. A. (1997). *Believing in magic: The psychology of superstition*. Oxford: Oxford University Press.
- Wellman, H. M. & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, 43, 337-375.
- White, P. A. (1995). *The understanding of causation and the production of action. From infancy to adulthood*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Wilkening, F. (1979). Combining of stimulus dimensions in children's and adults' judgments of area: An information integration analysis. *Developmental Psychology*, 15(1), 25-33.
- Wilkening, F. (1981). Integrating velocity, time, and distance information: A developmental study. *Cognitive Psychology*, 13, 231-247.

- Wilkening, F. (1982). Children's knowledge about time, distance, and velocity interrelations. In W. J. Friedman (Hrsg.), *The developmental psychology of time*. New York: Academic Press.
- Wilkening, F. (1989). Adaptives Denken von Kindern: Neue Aufgaben in der kognitiven Entwicklung. In W. Schönflug (Hrsg.), *Bericht über den 36. Kongress der Deutschen Gesellschaft für Psychologie, Berlin 1988*. Göttingen: Hogrefe.
- Wilkening, F. & Anderson, N. H. (1991). Representation and diagnosis of knowledge structures in developmental psychology. In N. H. Anderson (Hrsg.), *Contributions to information integration theory* (Bd. 3: Developmental, S. 45-80). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.
- Wilkening, F. & Lamsfuss, S. (1993). (Miss-)Konzepte der naiven Physik im Entwicklungsverlauf. In W. Hell, K. Fiedler & G. Gigerenzer (Hrsg.), *Kognitive Täuschungen: Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns* (S. 271-290). Heidelberg: Spektrum.
- Wilkening, F., Levin, I. & Druyan, S. (1987). Children's counting strategies for time quantification and integration. *Developmental Psychology*, 23(6), 823-831.
- Woolley, J. D. (1997). Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development*, 68(6), 991-1011.
- Woolley, J. D. & Phelps, K. E. (1994). Young children's practical reasoning about imagination. *British Journal of Developmental Psychology*, 12(1), 53-67.

Anhang A: Instruktion des Experimentes

Wichtige Instruktion für die Versuchsleiterin: Sie soll sich sowohl an die Reihenfolge der Erfassungsblätter als auch an die Reihenfolgen der Variablen und Antwortinstruktionen auf den Erfassungsblättern halten, da die unabhängigen und abhängigen Variablen, die Antwortinstruktionen sowie Antwortalternativen mehrerer Variablen in ihren Reihenfolge variiert wurden.

Vor Beginn des Experimentes werden Geschlechtszugehörigkeit, Geburtsdatum und Schulklasse der Versuchsperson sowie das Datum des Erhebungstages notiert. Bei jüngeren Kindern wird das Geburtsdatum durch die Lehrperson in Erfahrung gebracht.

1. Einleitung

Ich bin in einer Zauberschule und lerne zaubern. Wenn ich mit der Schule fertig bin, möchte ich ein Zauberprogramm für Kinder auf die Beine stellen. Bevor ich aber das Programm vor den Kindern auf einer Bühne vorführen werde, probiere ich verschiedene Zaubertricks zuerst bei verschieden altrigen Kindern aus. Ich möchte von den Kindern wissen: Sind meine Vorführungen überhaupt Zaubertricks für Kinder? Merken die Kinder, wann ich zaubere und wann nicht. Merken die Kinder also, ob es ein Trick ist oder ob es gar keiner ist? Wie gut finden die Kinder meine Tricks?

Heute habe ich den Farbmischungstrick mitgebracht. Wie du siehst, hat es in diesen Gläsern Flüssigkeiten mit verschieden hellem Rotviolett. Ich werde nachher je zwei verschiedene Flüssigkeiten miteinander vermischen, indem ich sie zusammengiesse. Beim Mischen werde ich nun aber zaubern. Das entscheidende Moment bei meinen Vorführungen ist also dann gegeben, wenn ich die Flüssigkeiten tatsächlich zusammengiesse. Manchmal zaubere ich dabei, manchmal mehr, manchmal weniger, vielleicht auch gar nicht. Du musst nun herausfinden, ob ich überhaupt gezaubert habe und wie fest ich gezaubert habe.

*Hast du noch fragen oder sollen wir beginnen?*²³

²³ In der vorliegenden Instruktion werden alle Anweisungen, die sich direkt an die Versuchspersonen richten, kursiv dargestellt.

2. Prüfung der Fähigkeit zur Helligkeitsdiskrimination und des Verständnisses von Seriation

Voraussetzungen des Experimentes, die erfüllt sein müssen, sind die Fähigkeit zur Helligkeitsunterscheidung der verwendeten gefärbten Flüssigkeiten sowie das Verständnis von Seriation anhand des Versuchsmaterials. Es muss die folgende Frage richtig beantwortet werden: Kann die Versuchsperson die vier verschiedenen Farbtintensitäten 1, 2, 3 und 4 unterscheiden? Hierzu soll die Versuchsperson die Gläser selbst nach ansteigender oder absteigender Helligkeit ordnen.

Ganz am Anfang möchte ich von dir wissen, ob du diese verschieden hellen Flüssigkeiten überhaupt unterscheiden kannst. Ich habe hierzu diese vier Gläser (die vier Gläser der Skala ungeordnet auf das Podest stellen) extra in Unordnung gebracht. Ich möchte, dass du diese wieder in die richtige Reihenfolge der Helligkeit nach bringst.

3. Erfassung der abhängigen Variablen pro Versuchsdurchgang vor der Präsentation des Mischungsvorganges

Am Anfang des Experimentes steht immer der Probeversuchsdurchgang, bei welchem aus der Mischung von zwei farblosen Flüssigkeiten eine rotviolette Flüssigkeit resultiert. Im Anschluss daran folgen die Versuchsdurchgänge, deren Anordnungen variieren. Die folgenden abhängigen Variablen werden in der hier dargelegten Reihenfolge erfasst.

A. Erfassung der prädiktiven Regeln

Bevor ich die Flüssigkeiten mische, möchte ich von dir wissen, was denn eigentlich beim Mischen passieren würde ohne Zaubern.

An dieser Stelle werden die beiden Gläser des Probeversuchsdurchganges auf das Podest gestellt.

Probeversuchsdurchgang

Wie du siehst, sind die Flüssigkeiten in beiden Gläsern farblos, es sieht wie Wasser aus. Angenommen, das da drin ist normales Wasser: Was würde passieren, wenn ich den Inhalt dieses Glases (eines der beiden Gläser in die Hand nehmen) in das andere Glas

leeren würde, ohne zu zaubern (Handbewegung des Kippens machen): Würde es heller werden? Würde es dunkler werden? Oder würde es gleich bleiben? Die Antwortalternativen wurden jeweils ausbalanciert.

Versuchsdurchgänge mit gleicher Intensitätskombination

Für die beiden Versuchsdurchgänge bei gleicher Intensitätskombination gilt folgende Anleitung (jedes Mal werden die zwei Ursprungsgläser auf das Podest gestellt):

Hier sind die beiden Flüssigkeiten gleich hell. Ich werde nachher diese beiden Flüssigkeiten mischen. Angenommen, das da drin ist normaler Sirup: Was würde passieren, wenn ich diese Flüssigkeit (eines der beiden Gläser in die Hand nehmen) in das andere Glas leeren würde, ohne zu zaubern (Handbewegung des Kippens machen): Würde es heller werden? Würde es dunkler werden? Oder würde es gleich bleiben? Die Antwortalternativen werden jeweils ausbalanciert.

Versuchsdurchgänge mit unterschiedlicher Intensitätskombination

Für die beiden Versuchsdurchgänge bei unterschiedlicher Intensitätskombination gilt Folgendes (jedes Mal werden die zwei Ursprungsgläser auf das Podest gestellt):

Hier sind die beiden Flüssigkeiten nicht gleich hell. Ich werde nachher die hellere Flüssigkeit (Glas mit der helleren Flüssigkeit in die Hand nehmen) in die dunklere leeren (Handbewegung des Kippens machen). Was würde passieren, ohne zu zaubern: Würde es heller werden? Würde es dunkler werden? Oder würde es gleich bleiben? Die Antwortalternativen werden jeweils ausbalanciert.

B. Erfassung der subjektiven Sicherheiten der prädiktiven Regel vor Beobachtung des Mischungsvorganges

Ich möchte von dir wissen, wie sicher du dir bist, dass die Mischung ohne Zaubern ... (Antwort wiederholen, die vorher gegeben wurde) werden würde. Hierzu habe ich ein Blatt, auf dem du mir das anzeigen kannst (Ratingskala für die subjektive Sicherheitserfassung in der prädiktiven Bedingung hervorheben und vor die Versuchsperson legen). Wie du siehst, hat es hier Zahlen von null bis zehn (mit einem Stift auf die Zahlen zeigen) und zwei Männchen auf je einer Seite. Das Männchen hier (auf das linke Männchen zeigen) kratzt sich am Kopf und weiss nicht, ob es die richtige Antwort gegeben hat oder nicht. Wenn du also gar nicht sicher bist, ob es ohne Zaubern ... (nochmals Antwort wiederholen, die vorher gegeben wurde) würde, dann zeigst du mir hier mit

dem Finger drauf. Wenn du aber doch etwas sicher bist, kannst du weiter oben anzeigen (mit dem Stift auf der Skala von 0 zu 5 fahren). Hier ist die Mitte, d. h. wenn du also mittel sicher bist, kannst du hier anzeigen. Aber wenn du mehr als mittelsicher bist, kannst du auch weiter hinauf gehen (mit dem Stift auf der Skala von 5 zu 10 fahren) bis zum anderen Männchen, das ganz sicher ist, dass die Mischung ohne Zaubern ... (nochmals Antwort wiederholen, die vorher gegeben wurde) werden würde. Die Instruktion mit Beginn bei Skalenpunkt null wird jeweils abgewechselt mit der Instruktion mit Beginn bei Skalenpunkt zehn.

4. Erfassung der abhängigen Variablen pro Versuchsdurchgang nach der Präsentation des Mischungsvorganges

Jetzt wird es spannender, denn jetzt werde ich die Flüssigkeiten tatsächlich mischen.

Die Präsentationen der Mischungsvorgänge der einzelnen Versuchsdurchgänge werden in derselben Reihenfolge präsentiert, wie sie zuvor bei der Erfassung der abhängigen Variablen vor der Präsentation verwendet wurde. Am Anfang des Experiments steht also erneut der Probeversuchsdurchgang, gefolgt von den variierten Versuchsdurchgängen. Die folgenden abhängigen Variablen werden in der hier dargelegten Reihenfolge erfasst.

C. Erfassung der Realitätseinschätzungen der Mischungsergebnisse anhand der Zauberskala

Das entsprechende Gläserpaar wird wieder hervorgenommen und die Deckel werden entfernt.

Weisst du noch, was du vorher gesagt hast, was passieren würde ohne Zaubern? (Falls keine Erinnerung mehr oder eine falsche Antwort gegeben wird: Die frühere Antwort geben.) Du musst jetzt bei den farbigen Flüssigkeiten (gilt nicht für Probendurchgang) genau hinschauen, um zu sehen, was passiert. Damit es einfacher wird, stelle ich dir ein Vergleichsglas dazu (Referenzglas dazu stellen), dessen Flüssigkeit genau die gleiche Helligkeit hat, wie die dunklere Flüssigkeit der beiden anderen Gläser. Damit kannst du auch nachher noch vergleichen, was passiert ist. Aber schau jetzt trotzdem genau zu, was passiert. (Die Flüssigkeiten werden zusammengeleert; bei den unterschiedlichen Intensitäten wird die hellere Flüssigkeit in die dunklere geleert.) Was ist passiert? (Falls

keine Antwort kommt, variiert vorgeben: heller, dunkler, gleich geblieben?) Die Antwort wird erfasst. *Habe ich jetzt gezaubert und wenn ja, wie fest? Ich möchte, dass du mir das auf dieser Skala anzeigst* (Zauberskala hervorheben und vor die Versuchsperson legen). *Wenn du findest ich habe gar nicht gezaubert, dann zeigst du mir da an, wo es kein Zaubermännchen hat* (mit einem Stift auf die entsprechende Stelle zeigen und in der Folge die Instruktion mit Zeigen begleiten). *Wenn du findest, ich habe ein bisschen gezaubert, zeigst du mir da an, wo es wenige Zaubermännchen hat. Wenn du findest, ich habe mittelfest gezaubert, zeigst du mir die Mitte an. Wenn du der Meinung bist, ich habe ziemlich fest gezaubert, zeigst du mir da an, wo es viele Zaubermännchen hat. Wenn du findest, ich habe ganz fest gezaubert, zeigst du mir da an, wo es am meisten Zaubermännchen hat.* Die Instruktion mit Beginn bei Skalenpunkt null wird jeweils abgewechselt mit der Instruktion mit Beginn bei Skalenpunkt zehn.

Wenn Schwierigkeiten beim Begriff „Zaubern“ auftreten, wird eine Alternative geboten: *Wenn du denkst, das ist ja nur ein Trick, dann sollst du mir sagen, wie fest ich getrickst habe. (Oder „bschisse“).*

D. Erfassung der subjektiven Sicherheiten der prädiktiven Regel nach Beobachtung des Mischungsvorganges

Nun hast du vorher gesagt, es würde ohne Zaubern ... (Antwort wiederholen, die vorher in der prädiktiven Bedingung gegeben wurde) *werden. Würdest du nun, nachdem du meinen Trick gesehen hast, nochmals dasselbe sagen, wieder im Falle von ohne Zaubern, oder würdest du etwas anderes sagen? Hierzu habe ich wieder ein Blatt, auf dem du mir das anzeigen kannst* (Ratingskala für die subjektive Sicherheitserfassung nach Beobachtung des Mischungsvorganges hervorheben und vor die Versuchsperson legen). *Wie du siehst, hat es hier Zahlen von minus fünf bis plus fünf* (mit einem Stift auf die Zahlen zeigen) *und zwei Gesichter auf je einer Seite. Das Gesicht hier* (auf das linke Gesicht zeigen) *schaut grimmig drein, weil es im Nachhinein etwas anderes sagen würde, wenn nochmals gefragt würde, was ohne Zaubern passiert. Wenn du also nun denkst, du hättest vorher etwas anderes sagen sollen, als ich gefragt habe, was denn ohne Zaubern passieren würde, dann zeigst du mit dem Finger da drauf. Wenn du aber nicht so sicher bist, aber doch denkst, du hättest etwas anderes sagen sollen, dann zeigst du hier drauf* (mit dem Stift stets die Instruktion mit Zeigen begleiten). *Wenn du gar nicht mehr weisst, was du nun sagen würdest, zeigst du die Mitte an. Wenn du doch*

eher sicher bist, dass du nochmals das Gleiche antworten würdest, gehst du weiter auf die andere Seite. Hier beim zufriedenen Gesicht zeigst du drauf, wenn du nochmals genau das Gleiche sagen würdest. Die Instruktion mit Beginn bei Skalenpunkt minus fünf wird jeweils abgewechselt mit der Instruktion mit Beginn bei Skalenpunkt plus fünf.

E. Zuschreibung der Machbarkeit zu verschiedenen Akteuren

Jetzt würde mich noch interessieren, ob du meinst, du könntest diesen Trick (immer Trick sagen!) selber auch. Du kannst bei der Antwort auswählen zwischen ja, eher ja, vielleicht, eher nein oder nein (die Reihenfolge der Antwortalternativen wird jeweils alternierend umgedreht). Und was meinst du, könnten das deine Eltern auch? Könnten dein Mami oder dein Papi das auch? Gleiche Antwortalternativen mit Variation geben. Meinst du, Frau ... / Herr ... (Name der Lehrperson) könnte diesen Trick auch? Gleiche Antwortalternativen mit Variation geben. Was meinst du, könnte ein anderer Zauberer oder eine andere Zauberin das auch? Gleiche Antwortalternativen mit Variation geben.

Falls gefragt wird, ob sie dieselben Gläser mit den Flüssigkeiten zur Verfügung hätten, so wird dies verneint.

5. Mögliche Erklärungen zur Entstehung des Farbmischungstricks

Nun sind wir fertig. Ganz zum Schluss möchte ich aber noch wissen, ob du eine Idee oder eine Ahnung hast, wie ich das gemacht habe. Falls keine spontane Antwort kommt, soll direkt gefragt werden: *Glaubst du, dass es richtige Zauberkraft braucht, um diesen Zaubertrick zu können?* Die offenen Antworten sollen stichwortartig notiert werden.

6. Debriefing

Die Versuchsperson soll am Ende über das Geheimnis des Tricks (entgegen dem Usus bei Zauberern!) eingeweiht werden. Da aber zuerst alle Kinder derselben Gruppe am Experiment teilgenommen haben müssen, werden sie am Schluss der Versuchsdurchführung mit allen Teilnehmenden in der ganzen Gruppe aufgeklärt. Falls dies nicht möglich ist, wird die letzte Versuchsperson aufgeklärt, mit dem Auftrag, es den anderen Kindern weiter zu erzählen.

Die Versuchsleiterin bedankt sich für die Teilnahme am Experiment. Die Kinder bekommen eine Süßigkeit, die sie sich selbst auslesen dürfen.

Anhang B: Farbintensitätsskala



Darstellung der Farbintensitätsskala mit vier Helligkeitsstufen (links: Helligkeit 1, rechts: Helligkeit 4). Die Farbintensitätsskala diente der Überprüfung der Fähigkeiten zur Farbhelligkeitsdiskrimination und der Seriation im Vortest. Die Helligkeiten der Flüssigkeiten der verschiedenen Versuchsdurchgänge entsprachen den Helligkeiten der Skala von 1 bis 4.

Anhang C: Darstellung der einzelnen Versuchsdurchgänge

1. Probeversuchsdurchgang



Ausgangssituation vor dem Mischungsvorgang: Ursprungsgläser mit farblosen Flüssigkeiten.

Abhängige Variablen (Phase 1):

- prädiktive Mischungsregel
- Subjektive Sicherheit des prädiktiven Regel



Durchführung des Mischungsvorgangs:

- mit chemischer Manipulation
- vor den Augen der Versuchsperson



Physikalisch unkorrektes Mischungsergebnis:

Abhängige Variablen (Phase 2 und 3):

- Einschätzung des Realitätscharakters
- Einschätzung der subjektiven Sicherheit der prädiktiven Regel
- Machbarkeitszuschreibungen

2. Physikalisch korrektes Mischungsergebnis gleicher Ursprungsintensitäten



Ausgangssituation vor dem Mischungsvorgang: links Referenzglas; rechts: Ursprungsgläser mit gleicher Intensitätskombination.

Abhängige Variablen (Phase 1):

- prädiktive Mischungsregel
- Subjektive Sicherheit des prädiktiven Regel



Durchführung des Mischungsvorgangs:

- ohne chemische Manipulation
- vor den Augen der Versuchsperson



Physikalisch korrektes Mischungsergebnis:

Abhängige Variablen (Phase 2 und 3):

- Einschätzung des Realitätscharakters
- Einschätzung der subjektiven Sicherheit der prädiktiven Regel
- Machbarkeitszuschreibungen

3. Physikalisch unkorrekte Mischung gleicher Ursprungsintensitäten



Ausgangssituation vor dem Mischungsvorgang: links Referenzglas; rechts: Ursprungsgläser mit gleicher Intensitätskombination.

Abhängige Variablen (Phase 1):

- prädiktive Mischungsregel
- Subjektive Sicherheit des prädiktiven Regel



Durchführung des Mischungsvorgangs:

- mit chemischer Manipulation
- vor den Augen der Versuchsperson



Physikalisch unkorrektes Mischungsergebnis:

Abhängige Variablen (Phase 2 und 3):

- Einschätzung des Realitätscharakters
- Einschätzung der subjektiven Sicherheit der prädiktiven Regel
- Machbarkeitszuschreibungen

4. Physikalisch korrekte Mischung unterschiedlicher Ursprungsintensitäten



Ausgangssituation vor dem Mischungsvorgang: links Referenzglas; rechts: Ursprungsgläser mit unterschiedlicher Intensitätskombination.

Abhängige Variablen (Phase 1):

- prädiktive Mischungsregel
- Subjektive Sicherheit des prädiktiven Regel



Durchführung des Mischungsvorgangs:

- ohne chemische Manipulation
- vor den Augen der Versuchsperson



Physikalisch korrektes Mischungsergebnis:

Abhängige Variablen (Phase 2 und 3):

- Einschätzung des Realitätscharakters
- Einschätzung der subjektiven Sicherheit der prädiktiven Regel
- Machbarkeitszuschreibungen

5. Physikalisch unkorrekte Mischung unterschiedlicher Ursprungsintensitäten



Ausgangssituation vor dem Mischungsvorgang: links Referenzglas; rechts: Ursprungsgläser mit unterschiedlicher Intensitätskombination.

Abhängige Variablen (Phase 1):

- prädiktive Mischungsregel
- Subjektive Sicherheit des prädiktiven Regel



Durchführung des Mischungsvorgangs:

- mit chemischer Manipulation
- vor den Augen der Versuchsperson



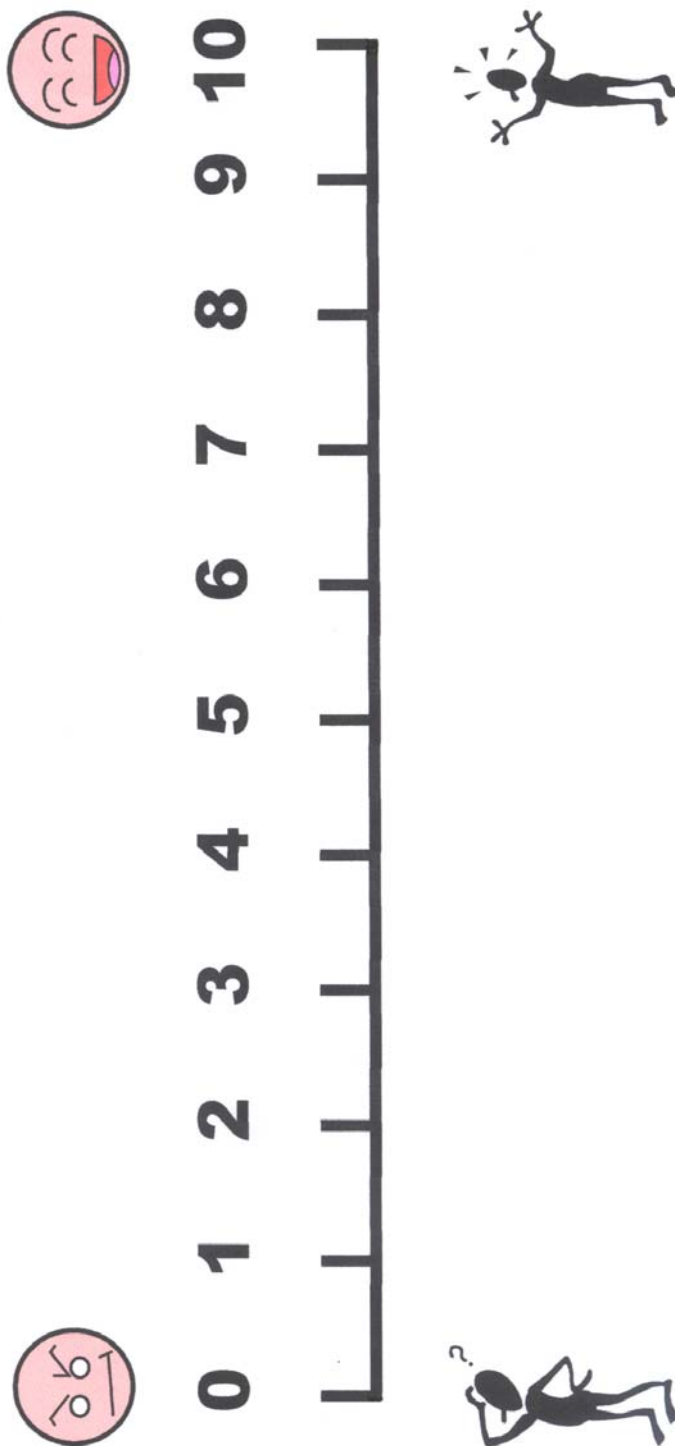
Physikalisch unkorrektes Mischungsergebnis:

Abhängige Variablen (Phase 2 und 3):

- Einschätzung des Realitätscharakters
- Einschätzung der subjektiven Sicherheit der prädiktiven Regel
- Machbarkeitszuschreibungen

Anhang D: Rating-Skalen

Ratingskala 1: Erfassung der subjektiven Sicherheit der prädiktiven Mischungsurteile

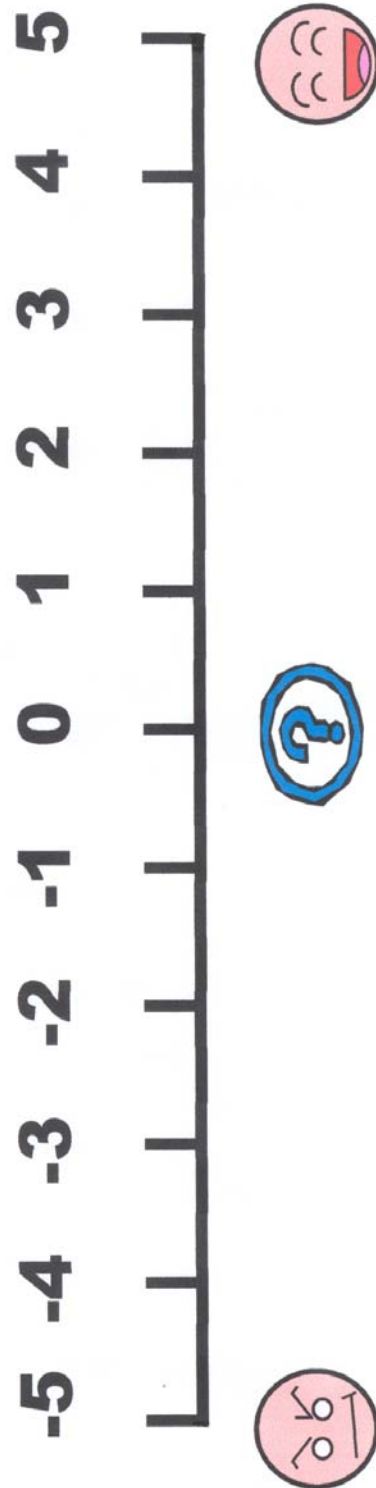


Skala zur Einschätzung der subjektiven Sicherheit des vorausgesagten Mischungsergebnisses (Phase 1): „Wie sicher bist Du, dass es ohne Zaubern ... (genanntes Mischungsurteil wiederholen) wird (0-10)?“

Ratingskala 2: Erfassung der Realitätseinschätzung der Mischungsresultate



Zauberskala zur Realitätseinschätzung der verschiedenen Mischungsresultate nach Beobachtung des Mischungsvorgangs (Phase 2): „Wie fest habe ich beim Mischungsvorgang gezaubert (0-10)?“

Ratingskala 3: Erfassung der subjektiven Regelsicherheit nach dem Mischungsvorgang

Skala zur Erfassung der subjektiven Sicherheit der prädiktiven Mischungsregel nach Beobachtung des Mischungsvorganges in Phase 2 (-5 bis +5): Verwerfung der ursprünglichen Regel im negativen Bereich; absolute Unsicherheit bzw. weder Verwerfung noch Bestätigung der ursprünglichen Regel bei Wert null; Bestätigung der ursprünglichen Regel im positiven Bereich.

Danksagung

An dieser Stelle möchte ich allen danken, die zur Verwirklichung der vorliegenden Arbeit beigetragen haben.

An erster Stelle bedanke ich mich bei Herrn Prof. Dr. F. Wilkening für die Themenstellung, die wissenschaftliche Förderung und Betreuung sowie für seine freundliche und ermutigende Unterstützung während der ganzen Projektphase. Herrn Prof. Dr. W. Marx sei herzlich für seine freundliche Bereitschaft zur Begutachtung meiner Dissertation gedankt.

Grosser Dank gilt den Kindern und Erwachsenen, welche an der Untersuchung teilgenommen haben. Den Studentinnen Patricia Schär, Silvia Balsama Erlbacher, Patrizia Belardo, Mirjam Lehner und Erika Stüssi danke ich für die Mithilfe bei der Datenerhebung. Bei meinem Teamkollegen Andreas Rapp bedanke ich mich herzlich für die wertvollen wissenschaftlichen Beiträge und Diskussionen.

Mein persönlicher Dank gilt meinem Ehemann Stefan Haller für seine liebevolle und förderliche Unterstützung während der ganzen Zeitspanne der Dissertation. Meinen Eltern, die mich während des Studiums stark unterstützt haben, möchte ich an dieser Stelle nochmals meinen herzlichen Dank aussprechen.

Zürich, im November 2004

Priska Schmid Haller

Curriculum vitae von Priska Simone Schmid Haller

Persönliche Angaben

Name	Priska Simone Schmid Haller
Geburtsdatum	26. Mai 1970
Bürgerort	Buchs ZH, Zürich und Reinach AG

Ausbildung

1977 – 1983	Primarschule, Bachenbülach
1983 – 1989	Gymnasium, Kantonsschule Zürcher Unterland, Bülach, Maturität Typus B
1990 – 1994	Studium der Humanmedizin, Universität Zürich
1994 – 2001	Studium der Psychologie, Universität Zürich (lic. phil.) Hauptfach: Psychologie, Vertiefungsrichtung Klinische Psychologie Erstes Nebenfach: Psychopathologie des Erwachsenenalters Zweites Nebenfach: Allgemeine Religionsgeschichte Titel der Lizarbeit: Rationalität und Esoterik – Die Rolle der Esoterik im Lebensentwurf von Akademikern und Akademikerinnen (qualitative Pilotstudie im 1. NF bei Prof. Dr. C. Scharfetter)
seit Oktober 2001	Assistentin und Doktorandin in der Fachgruppe Allgemeine und Entwicklungspsychologie, Prof. Dr. Friedrich Wilkening, Psychologisches Institut, Universität Zürich

